

Fr m: Chan, Christina
Sent: Monday, February 10, 2003 4:01 PM
To: Davis, Minh-Tam; STIC-Biotech/ChemLib
Subject: RE: Rush search request for 09/939293

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: Davis, Minh-Tam
Sent: Monday, February 10, 2003 3:48 PM
To: Chan, Christina
Subject: Rush search request for 09/939293

Please search in commercial database, PGPUB and issued patent files:

- 1) Residues 56-139, 56-239 of SEQ ID NO:1, with and without size limitation to the size of the corresponding sequences.
- 2) SEQ ID NO:13, with and without size limitation to the size of the corresponding sequence.
- 3) The amino acid sequence comprising the sequence Ala-Val of residues 56-139, 56-239 of SEQ ID NO:1.

Thank you.

MINH TAM DAVIS

ART Unit 1642, room 8A01, MB 8E12
305-2008

08/2000

RECEIVED
FEB 10 2003
STIC-BIOTECH/CHM LIB
(STIC)

Point of Contact:
Thomas G. Larson, Ph.D.
703-308-7309
CM1, Rm. 6 B 01

Searcher: Larson
Phone: _____
Location: _____
Date Picked Up: 2/11
Date Completed: 2/19
Searcher Prep/Review: 20
Clerical: _____
Online time: 10

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 42
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: ABSS05
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Larson, Thom

87004

From: Davis, Minh-Tam
Sent: Thursday, February 20, 2003 10:50 AM
To: Larson, Thom
Subject: RE: Search for 09/939293

rcc.
2/20/03
TOL

Yes. Thanks

-----Original Message-----

Fr m: Larson, Thom
Sent: Thursday, February 20, 2003 10:31 AM
T : Davis, Minh-Tam
Subject: RE: Search for 09/939293

Do you need a rush search?

-----Original Message-----

From: Davis, Minh-Tam
Sent: Thursday, February 20, 2003 10:29 AM
To: Larson, Thom
Subject: Search for 09/939293

Could you do an additional search for this case?
Please search in commercial database, PG PUB, and issued patent files:
1) the amino acid sequence encoded by SEQ ID NO:1.
Thank you.
MINH TAM DAVIS
ART UNIT 1642, ROOM 8A01, MB 8E12
305-2008

Point of Contact:
Thomas G. Larson, Ph.D.
703-308-7309
CM1, Rm. 6 B 01

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 02:29:44 ; Search time 13807 Seconds

(without alignments)
177.057 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Sequence: 1 actcaatctctcaggtacag.....gtctctcagaattgataaga 84

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_com:*
21: em_or:*
22: em_ov:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vl:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 84 | 100.0 | 227 | 6 | AX069562 Sequence |
| 2 | 84 | 100.0 | 720 | 6 | AR107983 Sequence |
| 3 | 84 | 100.0 | 720 | 6 | AX191368 Sequence |
| 4 | 84 | 100.0 | 1358 | 6 | AX392556 Sequence |
| 5 | 84 | 100.0 | 1358 | 6 | AP262240 Sequence |
| 6 | 84 | 100.0 | 1371 | 9 | BC011509 Homo sapi |
| 7 | 84 | 100.0 | 1374 | 9 | AK024768 Homo sapi |
| 8 | 82.4 | 98.1 | 1327 | 9 | AK057778 Homo sapi |
| 9 | 77.6 | 92.4 | 720 | 6 | AR107984 Sequence |
| 10 | 77.6 | 92.4 | 720 | 6 | AR107985 Sequence |
| 11 | 77.6 | 92.4 | 720 | 6 | AR107986 Sequence |
| 12 | 77.6 | 92.4 | 720 | 6 | AX191370 Sequence |
| 13 | 77.6 | 92.4 | 720 | 6 | AX191371 Sequence |
| 14 | 77.6 | 92.4 | 720 | 6 | AX191372 Sequence |
| 15 | 76 | 90.5 | 189947 | 2 | AC048338 Homo sapi |
| 16 | 76 | 90.5 | 323263 | 3 | AC079406 Homo sapi |
| 17 | 71.2 | 84.8 | 720 | 6 | AR107987 Sequence |
| 18 | 71.2 | 84.8 | 720 | 6 | AR107988 Sequence |
| 19 | 71.2 | 84.8 | 720 | 6 | AR107989 Sequence |
| 20 | 71.2 | 84.8 | 720 | 6 | AX191373 Sequence |
| 21 | 71.2 | 84.8 | 720 | 6 | AX191374 Sequence |
| 22 | 71.2 | 84.8 | 720 | 6 | AX191375 Sequence |
| 23 | 64.8 | 77.1 | 432 | 6 | AX070289 Sequence |
| 24 | 54 | 64.3 | 376 | 6 | AX071845 Sequence |
| 25 | 50.2 | 59.8 | 383 | 6 | AX071832 Sequence |
| 26 | 44.4 | 52.9 | 1356 | 10 | AF203914 Mus muscu |
| 27 | 44.4 | 52.9 | 2585 | 10 | BC024780 Mus muscu |
| 28 | 43.4 | 51.7 | 224847 | 2 | AC129569 Mus muscu |
| 29 | 30.2 | 36.0 | 161862 | 2 | AC096077 Rattus no |
| 30 | 30.2 | 36.0 | 234855 | 2 | AC068843 Homo sapi |
| 31 | 29.8 | 35.5 | 52831 | 2 | AC100274 Mus muscu |
| 32 | 29.8 | 35.5 | 73546 | 2 | AC129938 Mus muscu |
| 33 | 29.8 | 35.5 | 171595 | 2 | AL391476 Human DNA |
| 34 | 29.8 | 35.5 | 201815 | 2 | AC113497 Mus muscu |
| 35 | 29.6 | 35.2 | 99798 | 2 | AC117321 Rattus no |
| 36 | 29.6 | 35.2 | 106359 | 2 | AC103443 Rattus no |
| 37 | 29.6 | 35.2 | 162451 | 9 | AL353597 Human DNA |
| 38 | 29.6 | 35.2 | 184469 | 2 | AL365447 Homo sapi |
| 39 | 29.6 | 35.2 | 204153 | 9 | AC074091 Homo sapi |
| 40 | 29.2 | 34.8 | 161881 | 2 | AC021685 Homo sapi |
| 41 | 29.2 | 34.8 | 163798 | 2 | AC012127 Homo sapi |
| 42 | 29.2 | 34.8 | 168901 | 2 | AC079598 Homo sapi |
| 43 | 29.2 | 34.8 | 187898 | 2 | AC130851 Rattus no |
| 44 | 29.2 | 34.8 | 194481 | 2 | AC126068 Rattus no |
| 45 | 29 | 34.5 | 124151 | 9 | AC008951 Homo sapi |

ALIGNMENTS

| RESULT 1 | AX069562 | 227 bp | DNA | linear | PAT 25-JAN-2001 |
|------------|---|-------------|-----|--------|-----------------|
| LOCUS | AX069562 | | | | |
| DEFINITION | Sequence 34 from Patent WO0102568. | | | | |
| ACCESSION | AX069562 | | | | |
| VERSION | AX069562.1 | GI:12579347 | | | |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | human. | | | | |
| | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 227) | | | | |
| AUTHORS | Williams,L.T., Escobedo,J., Inis,M.A., Garcia,P.D., Klinger,J., | | | | |
| | Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D., | | | | |
| | Lamson,G., Drmanac,R., Ckenjakov,R., Drmanac,S., Dickson,M., | | | | |

| Db | 116 | CGGTGTTCTCAGATTGATAAGA | 139 |
|---|---|------------------------|-----------------------------|
| RESULT 5 | AF262240 | 1358 bp | linear PRI 26-JUL-2000 |
| LOCUS | Homo sapiens Smac mRNA, complete cds; nuclear gene for mitochondrial product. | | |
| DEFINITION | AF262240 | | |
| ACCESSION | AF262240.1 | GI:9454218 | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | Homo sapiens. | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| REFERENCE | 1 (bases 1 to 1358) | | |
| AUTHORS | Du, C., Fang, M., Li, Y., Li, L. and Wang, X. | | |
| TITLE | Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition | | |
| JOURNAL | Cell 102 (1), 33-42 (2000) | | |
| MEDLINE | 20383536 | | |
| PUBMED | 10929711 | | |
| REFERENCE | 2 (bases 1 to 1358) | | |
| AUTHORS | Du, C., Fang, M., Li, Y. and Wang, X. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (01-MAY-2000) Howard Hughes Medical Institute and Department of Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235, USA | | |
| FEATURES | location/Qualifiers | | |
| SOURCE | 1..1358 | | |
| CDS | /organism="Homo sapiens" | | |
| | /db_xref="taxon:9606" | | |
| | 20..739 | | |
| | /function="binds IAPs and neutralizes their inhibition on caspase activation and activity" | | |
| | /note="antagonist of IAPs (inhibitors of apoptosis)" | | |
| | /codon_start=1 | | |
| | /product="Smac" | | |
| | /protein_id="AAE87716.1" | | |
| | /db_xref="GI:9454218" | | |
| | /translation="MAALKSMLSRVSPFRRCGLCPVAVANFKRCESELIRPMKR TVTSGFVTLCAVP IAKKSEPHSLSEALKRRAVLTDSSTPLSOTYALALAIIE YTRAVYTLTLYGQYSLILKRNSEEDDEVMOVIIIGARMTSHOEYLIETTMRA VGSSEMAERAYIOTGADQASITARNHILQVLQVEVHQLSRKETLAEHQIEELKQ KTGSEGERAESFEQEAFLRND" | | |
| | transit_peptide 20..184 | | |
| | /note="mitochondrial targeting sequence" | | |
| BASE COUNT | 341 a 314 c 358 g 345 t | | |
| ORIGIN | | | |
| Query Match | 100.0%; Score 84; DB 9; Length 1358; | | |
| Best Local Similarity | 100.0%; Pred. No. 1e-16; | | |
| Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| OY | 1 ACTTCATCTCTCAGGTACAGACAGGTGTGTGTCTCTGTGGGCTAACTTTAAGAG 60 | | |
| Db | 56 ACTTCATCTCTCAGGTACAGACAGGTGTGTGTCTCTGTGTGTGCTTAAGAG 115 | | |
| OY | 61 CGGTGTTCTCAGATTGATAAGA 84 | | |
| Db | 116 CGGTGTTCTCAGATTGATAAGA 139 | | |
| RESULT 6 | BC011909 | 1371 bp | mRNA linear PRI 02-AUG-2001 |
| LOCUS | Homo sapiens, similar to second mitochondrial-derived activator of caspase, clone MGC:19663 IMAGE:4137792, mRNA, complete cds. | | |
| ACCESSION | BC011909 | | |
| VERSION | BC011909.1 | GI:15080296 | |
| KEYWORDS | MGC. | | |
| SOURCE | Homo sapiens. | | |
| ORGANISM | Homo sapiens. | | |

| REFERENCE | AUTHORS | TITLE | JOURNAL | REMARK | COMMENT |
|-----------|-------------------|---|---------|--------|---------|
| 1 | (bases 1 to 1371) | Emkaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| 1 | (bases 1 to 1371) | Strausberg, R. | | | |
| 1 | (bases 1 to 1371) | Submitted (30-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | | | |
| 1 | (bases 1 to 1371) | NIH-MGC Project URL: http://mgc.nci.nih.gov | | | |
| 1 | (bases 1 to 1371) | Contact: MGC help desk | | | |
| 1 | (bases 1 to 1371) | Email: cgabs-remail.nih.gov | | | |
| 1 | (bases 1 to 1371) | Tissue Procurement: ATCC | | | |
| 1 | (bases 1 to 1371) | cDNA Library Preparation: Rubin Laboratory | | | |
| 1 | (bases 1 to 1371) | cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) | | | |
| 1 | (bases 1 to 1371) | DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), | | | |
| 1 | (bases 1 to 1371) | Galtherdsburg, Maryland; | | | |
| 1 | (bases 1 to 1371) | Web site: http://www.nisc.nih.gov/ | | | |
| 1 | (bases 1 to 1371) | http://www.nisc.nih.gov/ | | | |
| 1 | (bases 1 to 1371) | Shvachko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P. J., Tlionson, E.E., Touchman, J.W., Tsurgon, C., Vogt, D.L., Walker, M.A., Zhang, L., H. and Green, E.D. | | | |
| 1 | (bases 1 to 1371) | Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov | | | |
| 1 | (bases 1 to 1371) | Series: IRAL Plate: 28 Row: K Column: 12 | | | |
| 1 | (bases 1 to 1371) | This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10437143. | | | |
| 1 | (bases 1 to 1371) | Location/Qualifiers | | | |
| 1 | (bases 1 to 1371) | 1. 1371 | | | |
| 1 | (bases 1 to 1371) | /organism="Homo sapiens" | | | |
| 1 | (bases 1 to 1371) | /db_xref="taxon:9606" | | | |
| 1 | (bases 1 to 1371) | /clone="MGC:19863 IMAGE:4137792" | | | |
| 1 | (bases 1 to 1371) | /tissue_type="Muscle, Thadomyosarcoma" | | | |
| 1 | (bases 1 to 1371) | /clone_id="N1H_MGC_17" | | | |
| 1 | (bases 1 to 1371) | /lab_host="DH10B-R" | | | |
| 1 | (bases 1 to 1371) | /note="Vector: pOTB7" | | | |
| 1 | (bases 1 to 1371) | 33. 752 | | | |
| 1 | (bases 1 to 1371) | /ccoon_start=1 | | | |
| 1 | (bases 1 to 1371) | /product="Similar to second mitochondria-derived activator of caspase" | | | |
| 1 | (bases 1 to 1371) | /protein_id="AAH11909.1" | | | |
| 1 | (bases 1 to 1371) | /db_xref="GI:15080297" | | | |
| 1 | (bases 1 to 1371) | /translation="MALKSMISRSYSPFRYRCICLPVYANKRKFSELRIPMKH TVIIGFVTCICAPVIAKQSEPHSISSEALMRAYSIVTDSITPSTYALIAITE YTAAYVTLTSLRYOYTLILKMNSEEDVWQVIGARAEMTSKHORELLETTWTA VGSLEMALEAAYOGADQASITANNH1QLVKLOVEVHOLSRRKETLADROIELRO KTOEGEGERASEDEAVLYRED" | | | |
| 1 | (bases 1 to 1371) | BASE COUNT 349 a 314 c 363 g 345 t | | | |
| 1 | (bases 1 to 1371) | ORIGIN | | | |
| 1 | (bases 1 to 1371) | Query Match 100.0%; Score 84; DB 9; Length 1371; | | | |
| 1 | (bases 1 to 1371) | Best Local Similarity 100.0%; Prid. No. 1e-16; | | | |
| 1 | (bases 1 to 1371) | Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| 1 | (bases 1 to 1371) | QY 1 ACTCATCTCTTCAGGTACAGACAGTGTGTGTCGTCTGCTGTTGGCACTTAAGANG 60 | | | |
| 1 | (bases 1 to 1371) | DB 69 ACTTCATCTCTTCAGGTACAGACAGTGTGTGTCGTCTGCTGTTGGCACTTAAGANG 128 | | | |
| 1 | (bases 1 to 1371) | QY 61 CGGTGTTCTCAGAAATGATAGA 84 | | | |
| 1 | (bases 1 to 1371) | DB 129 CGGTGTTCTCAGAAATGATAGA 152 | | | |
| 1 | (bases 1 to 1371) | RESULT 7 | | | |
| 1 | (bases 1 to 1371) | AK024768 1374 bp mRNA linear PRI 29-SEP-2000 | | | |
| 1 | (bases 1 to 1371) | LOCUS | | | |
| 1 | (bases 1 to 1371) | DEFINITION Homo sapiens cDNA: FLJ21115 fis, clone CAS05491. | | | |

ACCESSION AK024768
VERSION GI:10437143
KEYWORDS oligo capping; fls (full insert sequence);
SOURCE Homo sapiens primary smooth muscle cells of human coronary artery
cDNA to mRNA, clone_lib:CAS clone:CAS05491.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (sites)
Matanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1374)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'-3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source location/Qualifiers
1. .1374
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CAS05491"
/cell_type="primary smooth muscle cells of human coronary
artery"
/clone_lib="CAS"
/note="Cloning vector pME18SFL3"
28. .747
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB1494.1"
/db_xref="GI:10437144"
/translation="MAALKSWLSRSVTSFFRYRQCLCPVYVANKRCFSELRPMHR
TVTIGFVTLCAVPDAQVYVLTSLRYTSLGKMSSEEDVQVITIGARAEWTSK
YTKAVYTLISLROYTSLGKMSSEEDVQVITIGARAEWTSKHOEYKLETTMTA
VGLSPMAEAAYOTGADADASTARNHIOLVKLYEVEHOLSRAETKLAENQIEELRO
KTOEGEERASEDEAVLRD"
BASE COUNT 351 a 314 c 361 g 348 t
ORIGIN
Query Match 100.0%; Score 84; DB 9; Length 1374;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTTCATCTTCAGGTACAGACAGTGTGTCCTGTTGGCTACTTAAGAAG 60
|||||
DB 64 ACTTCATCTTCAGGTACAGACAGTGTGTCCTGTTGGCTACTTAAGAAG 60
|||||
QY 61 CGGTGTTCTCAGAAATTGATAAGA 84
|||||
DB 124 CGGTGTTCTCAGAAATTGATAAGA 147
|||||
RESULT 8
AK057778 1327 bp mRNA linear PRI 31-OCT-2001
LOCUS
DEFINITION Homo sapiens cDNA FLJ25049 fls; clone CBL04001, highly similar to
ACCESSION AK057778
VERSION AK057778.1 GI:16553713
KEYWORDS oligo capping; fls (full insert sequence).

SOURCE Homo sapiens cerebellum cDNA to mRNA, clone_lib:CBL clone:CBL04001.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1
Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M.,
Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and
Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1327)
Sugano,S. and Suzuki,Y.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing; Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing; RAB; clone selection for
full insert sequencing; RAB and Helix Research Institute.
FEATURES
source location/Qualifiers
1. .1327
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CBL04001"
/tissue_type="cerebellum"
/clone_lib="CBL"
/note="Cloning vector: pME18SFL3"
28. .615
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71568.1"
/db_xref="GI:16553714"
/translation="MAALKSWLSRSVTSFFRYRQCLCPVYVANKRCFSELRPMHR
TVTIGFVTLCAVPDAQVYVLTSLRYTSLGKMSSEEDVQVITIGARAEWTSK
HOEYKLETTMTAVAGLSMAEAAYOTGADADASTARNHIOLVKLYEVEHOLSRA
KTEKLAENQIEELROKTOEGEERASEDEAVLRD"
BASE COUNT 317 a 309 c 354 g 347 t
ORIGIN
Query Match 98.1%; Score 82.4; DB 9; Length 1327;
Best Local Similarity 98.8%; Pred. No. 3.4e-16;
Matches 83; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACTTCATCTTCAGGTACAGACAGTGTGTCCTGTTGGCTACTTAAGAAG 60
|||||
DB 64 ACTTCATCTTCAGGTACAGACAGTGTGTCCTGTTGGCTACTTAAGAAG 123
|||||
QY 61 CGGTGTTCTCAGAAATTGATAAGA 84
|||||
DB 124 CGGTGTTCTCAGAAATTGATAAGA 147
|||||
RESULT 9
AR107984 720 bp DNA linear PAT 14-FEB-2001
LOCUS
DEFINITION Sequence 3 from patent US 6110691.
ACCESSION AR107984
VERSION AR107984.1 GI:12823471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 720)
Wang,X. and Du,C.
TITLE Activators of caspases

JOURNAL Patent: US 6110691-A 3 29-AUG-2000;
FEATURES
source 1..720
/organism="unknown"
BASE COUNT 207 a 145 c 190 g 178 t
ORIGIN

Query Match
Best Local Similarity 92.4%; Score 77.6; DB 6; Length 720;
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTGCTTCCTGTTGGCTACTTAAGAAG 60
|||||
Db 37 ACTTCATTTCTTCAGTACAGACAGTGTGTTGCTTCCTGTTGGCTACTTAAGAAG 96
|||||

QY 61 CGGTGTTTCTCAGATTTGATAGA 84
|||||
Db 97 CGGTGTTTCTCAGAAATGATAGA 120
|||||

RESULT 10
AR107985
LOCUS AR107985 720 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 4 from patent US 6110691.
ACCESSION AR107985
VERSION AR107985.1 GI:12823472
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang, X. and Du, C.
TITLE Activators of caspases
JOURNAL Patent: US 6110691-A 4 29-AUG-2000;
FEATURES
source 1..720
/organism="unknown"

BASE COUNT 202 a 154 c 192 g 172 t
ORIGIN

Query Match
Best Local Similarity 92.4%; Score 77.6; DB 6; Length 720;
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTGCTTCCTGTTGGCTACTTAAGAAG 60
|||||
Db 37 ACTTCATTTCTTCAGTACAGACAGTGTGTTGCTTCCTGTTGGCTACTTAAGAAG 96
|||||

QY 61 CGGTGTTTCTCAGATTTGATAGA 84
|||||
Db 97 CGGTGTTTCTCAGAAATGATAGA 120
|||||

RESULT 11
AR107986
LOCUS AR107986 720 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6110691.
ACCESSION AR107986
VERSION AR107986.1 GI:12823473
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang, X. and Du, C.
TITLE Activators of caspases
JOURNAL Patent: US 6110691-A 5 29-AUG-2000;
FEATURES
source 1..720
/organism="unknown"

BASE COUNT 205 a 151 c 188 g 176 t
ORIGIN

Query Match
Best Local Similarity 92.4%; Score 77.6; DB 6; Length 720;
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTGCTTCCTGTTGGCTACTTAAGAAG 60
|||||
Db 37 ACTTCATTTCTTCAGTACAGACAGTGTGTTGCTTCCTGTTGGCTACTTAAGAAG 96
|||||

QY 61 CGGTGTTTCTCAGATTTGATAGA 84
|||||
Db 97 CGGTGTTTCTCAGAAATGATAGA 120
|||||

RESULT 12
AX191370
LOCUS AX191370 720 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 3 from Patent WO0149719.
ACCESSION AX191370
VERSION AX191370.1 GI:15209590
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang, X. U. and Du, C. U.
TITLE Activators of caspases
JOURNAL Patent: WO 0149719-A 3 12-JUL-2001;
FEATURES
source 1..720
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Sequence"

BASE COUNT 207 a 145 c 190 g 178 t
ORIGIN

Query Match
Best Local Similarity 92.4%; Score 77.6; DB 6; Length 720;
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTGCTTCCTGTTGGCTACTTAAGAAG 60
|||||
Db 37 ACTTCATTTCTTCAGTACAGACAGTGTGTTGCTTCCTGTTGGCTACTTAAGAAG 96
|||||

QY 61 CGGTGTTTCTCAGATTTGATAGA 84
|||||
Db 97 CGGTGTTTCTCAGAAATGATAGA 120
|||||

RESULT 13
AX191371
LOCUS AX191371 720 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 4 from Patent WO0149719.
ACCESSION AX191371
VERSION AX191371.1 GI:15209591
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang, X. U. and Du, C. U.
TITLE Activators of caspases
JOURNAL Patent: WO 0149719-A 4 12-JUL-2001;
FEATURES
source 1..720
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Sequence"

BASE COUNT 202 a 154 c 192 g 172 t
ORIGIN

Query Match
Best Local Similarity 92.4%; Score 77.6; DB 6; Length 720;


```

* be preserved.
* 1 5226: contig of 5226 bp in length
* 5227 5326: gap of unknown length
* 5327 26441: contig of 21115 bp in length
* 26442 26541: gap of unknown length
* 26542 51092: contig of 24551 bp in length
* 51093 51192: gap of unknown length
* 51193 88442: contig of 37250 bp in length
* 88443 88542: gap of unknown length
* 88543 136963: contig of 48421 bp in length
* 136964 137064: gap of unknown length
* 137064 189947: contig of 52884 bp in length.

```

```

FEATURES
  source          1..189947
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="12"
                  /clone="RP11-512M8"

```

```

BASE COUNT      51804 a 44347 c 43777 g 49515 t    504 others
ORIGIN

```

```

Query Match      90.5%; Score 76; DB 2; Length 189947;
Best Local Similarity 94.0%; Pred. No. 1.6e-14;
Matches 79; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      1  ACTTCATTCTTCAGATACAGACAGTGTGTGTCGCTGTGCTACTTTAAGAAG 60
          |||||  ||||||||||||||||||||||||||||||||||||||||
Db 93120 ACTTCTACTTCAGATACAGACAGTGTGTGTCGCTGTGCTACTTTAAGAAG 93061

```

```

QY      61  CGGTGTTTCTCAGATATGATAGA 84
          ||||||||||||||||||||
Db 93060 CGGTGTTTCTCAGATATGATAGA 93037

```

```

Search completed: February 16, 2003, 00:27:27
Job time : 13849 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 12:12:34 : Search time 29.1493 Seconds
(without alignments)
883.757 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84
Sequence: 1 actctctctcagtcagtaag.....gttcctcgaattgataaga 84

Scoring table:
IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCUTUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 84 | 100.0 | 720 | 3 | US-09-479-309-1 |
| 2 | 77.6 | 92.4 | 720 | 3 | US-09-479-309-3 |
| 3 | 77.6 | 92.4 | 720 | 3 | US-09-479-309-4 |
| 4 | 77.6 | 92.4 | 720 | 3 | US-09-479-309-5 |
| 5 | 71.2 | 84.8 | 720 | 3 | US-09-479-309-6 |
| 6 | 71.2 | 84.8 | 720 | 3 | US-09-479-309-7 |
| 7 | 71.2 | 84.8 | 720 | 3 | US-09-479-309-8 |
| 8 | 24.6 | 29.3 | 12284 | 2 | US-08-937-102-1 |
| 9 | 24.6 | 29.3 | 12284 | 2 | US-08-876-991-1 |
| 10 | 24.6 | 29.3 | 12284 | 2 | US-09-059-853-1 |
| 11 | 24.2 | 28.8 | 532 | 4 | US-09-130-338-8 |
| 12 | 24.2 | 28.8 | 1968 | 4 | US-08-961-527-119 |
| 13 | 23.8 | 28.3 | 1636 | 4 | US-09-039-198A-1 |
| 14 | 23.8 | 28.3 | 1636 | 4 | US-08-877-599-1 |
| 15 | 23.8 | 28.3 | 1636 | 4 | US-09-267-574-1 |
| 16 | 23.8 | 28.3 | 1643 | 2 | US-08-486-839-3 |
| 17 | 23.8 | 28.3 | 1643 | 2 | US-08-486-839-5 |
| 18 | 23.8 | 28.3 | 1643 | 4 | US-09-151-011-3 |
| 19 | 23.8 | 28.3 | 1643 | 4 | US-09-343-623-3 |
| 20 | 23.8 | 28.3 | 1656 | 4 | US-09-039-198A-3 |
| 21 | 23.8 | 28.3 | 1656 | 4 | US-08-877-599-3 |
| 22 | 23.8 | 28.3 | 1713 | 2 | US-09-267-574-3 |
| 23 | 23.8 | 28.3 | 1713 | 2 | US-08-486-839-5 |
| 24 | 23.8 | 28.3 | 1713 | 3 | US-09-151-011-5 |
| 25 | 23.6 | 28.1 | 9707 | 4 | US-09-343-623-5 |
| 26 | 23.4 | 27.9 | 2244 | 4 | US-08-961-527-164 |
| 27 | 23.4 | 27.9 | 36159 | 4 | US-09-749-588-3 |

| | | | | | | | |
|---|----|------|------|--------|---|-------------------|--------------------|
| c | 28 | 23.4 | 27.9 | 168575 | 4 | US-09-426-290-1 | Sequence 1, Appl1 |
| | 29 | 23.2 | 27.6 | 1464 | 4 | US-09-004-838-126 | Sequence 126, App |
| | 30 | 23.2 | 27.6 | 21234 | 4 | US-09-810-671-3 | Sequence 3, Appl1 |
| c | 31 | 23 | 27.4 | 1262 | 4 | US-09-149-476-266 | Sequence 266, App |
| c | 32 | 23 | 27.4 | 1313 | 4 | US-09-149-476-112 | Sequence 112, App |
| c | 33 | 23 | 27.4 | 3742 | 2 | US-08-960-022-9 | Sequence 9, Appl1 |
| c | 34 | 23 | 27.4 | 12141 | 4 | US-09-488-671-10 | Sequence 10, Appl1 |
| c | 35 | 22.8 | 27.1 | 3393 | 4 | US-09-104-324B-1 | Sequence 1, Appl1 |
| c | 36 | 22.8 | 27.1 | 3393 | 4 | US-09-162-713-1 | Sequence 1, Appl1 |
| | 37 | 22.8 | 27.1 | 4698 | 1 | US-07-807-043B-5 | Sequence 5, Appl1 |
| | 38 | 22.8 | 27.1 | 4698 | 1 | US-08-299-849B-5 | Sequence 5, Appl1 |
| | 39 | 22.8 | 27.1 | 4698 | 2 | US-08-142-368A-5 | Sequence 5, Appl1 |
| | 40 | 22.8 | 27.1 | 4698 | 3 | US-08-967-727-5 | Sequence 5, Appl1 |
| c | 41 | 22.8 | 27.1 | 4698 | 4 | US-08-037-230D-5 | Sequence 5, Appl1 |
| c | 42 | 22.8 | 27.1 | 12616 | 2 | US-08-822-445-9 | Sequence 9, Appl1 |
| c | 43 | 22.8 | 27.1 | 12616 | 4 | US-09-396-540-9 | Sequence 9, Appl1 |
| c | 44 | 22.8 | 27.1 | 40328 | 3 | US-08-742-185-102 | Sequence 102, App |
| c | 45 | 22.6 | 26.9 | 693 | 4 | US-09-712-016-12 | Sequence 12, Appl1 |

ALIGNMENTS

RESULT 1
US-09-479-309-1
; Sequence 1, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTS00630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 720
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(717)
US-09-479-309-1

Query Match 100.0%; Score 84; DB 3; Length 720;
Best Local Similarity 100.0%; Pred. No. 3.4e-22;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCATTTCTGAGTACAGAGAGTGTGTGCTGTGCTGCTACTTAAGAG 60
DB 37 ACTCATTTCTGAGTACAGAGAGTGTGTGCTGTGCTGCTACTTAAGAG 96
QY 61 CGGTGTTCTCAGAAATGATAGA 84
DB 97 CGGTGTTCTCAGAAATGATAGA 120

RESULT 2
US-09-479-309-3
; Sequence 3, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTS00630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 720

```

: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-3
```

```
Query Match          92.4%; Score 77.6; DB 3; Length 720;
Best Local Similarity 95.2%; Pred. No. 8.1e-20;
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
Oy 1 ACTTCATCTTCAGGTACAGACAGGTTGTGTCTCTGTTGGCTAACTTTAAGAAG 60
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 37 ACTTCATCTTCAGGTACAGACAGGTTGTGTCTCTGTTGGCTAACTTTAAGAAG 96
Oy 61 CGGTGTTCTCAGAAATGATAGA 84
    ||||||| ||||||| |||||||
Db 97 CGGTGTTCTCAGAAATGATAGA 120
```

```

RESULT 3
US-09-479-309-4
: Sequence 4, Application US/09479309
: Patent No. 6110691
: GENERAL INFORMATION:
: APPLICANT: Wang, Xiaodong
: APPLICANT: Du, Chunying
: TITLE OF INVENTION: Activators of Caspases
: FILE REFERENCE: UTS0630
: CURRENT APPLICATION NUMBER: US/09/479,309
: CURRENT FILING DATE: 2000-01-06
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 4
: LENGTH: 720
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-4
```

```
Query Match          92.4%; Score 77.6; DB 3; Length 720;
Best Local Similarity 95.2%; Pred. No. 8.1e-20;
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
Oy 1 ACTTCATCTTCAGGTACAGACAGGTTGTGTCTCTGTTGGCTAACTTTAAGAAG 60
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 37 ACTTCATCTTCAGGTACAGACAGGTTGTGTCTCTGTTGGCTAACTTTAAGAAG 96
Oy 61 CGGTGTTCTCAGAAATGATAGA 84
    ||||||| ||||||| |||||||
Db 97 CGGTGTTCTCAGAAATGATAGA 120
```

```

RESULT 4
US-09-479-309-5
: Sequence 5, Application US/09479309
: Patent No. 6110691
: GENERAL INFORMATION:
: APPLICANT: Wang, Xiaodong
: APPLICANT: Du, Chunying
: TITLE OF INVENTION: Activators of Caspases
: FILE REFERENCE: UTS0630
: CURRENT APPLICATION NUMBER: US/09/479,309
: CURRENT FILING DATE: 2000-01-06
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 5
: LENGTH: 720
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
```

```

: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
: OTHER INFORMATION: Sequence
US-09-479-309-5
```

```
Query Match          92.4%; Score 77.6; DB 3; Length 720;
Best Local Similarity 95.2%; Pred. No. 8.1e-20;
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
Oy 1 ACTTCATCTTCAGGTACAGACAGGTTGTGTCTCTGTTGGCTAACTTTAAGAAG 60
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 37 ACTTCATCTTCAGGTACAGACAGGTTGTGTCTCTGTTGGCTAACTTTAAGAAG 96
Oy 61 CGGTGTTCTCAGAAATGATAGA 84
    ||||||| ||||||| |||||||
Db 97 CGGTGTTCTCAGAAATGATAGA 120
```

```

RESULT 5
US-09-479-309-6
: Sequence 6, Application US/09479309
: Patent No. 6110691
: GENERAL INFORMATION:
: APPLICANT: Wang, Xiaodong
: APPLICANT: Du, Chunying
: TITLE OF INVENTION: Activators of Caspases
: FILE REFERENCE: UTS0630
: CURRENT APPLICATION NUMBER: US/09/479,309
: CURRENT FILING DATE: 2000-01-06
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 6
: LENGTH: 720
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-6
```

```
Query Match          84.8%; Score 71.2; DB 3; Length 720;
Best Local Similarity 90.5%; Pred. No. 1.9e-17;
Matches 76; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
Oy 1 ACTTCATCTTCAGGTACAGACAGGTTGTGTCTCTGTTGGCTAACTTTAAGAAG 60
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 37 ACTTCATCTTCAGGTACAGACAGGTTGTGTCTCTGTTGGCTAACTTTAAGAAG 96
Oy 61 CGGTGTTCTCAGAAATGATAGA 84
    ||||||| ||||||| |||||||
Db 97 CGGTGTTCTCAGAAATGATAGA 120
```

```

RESULT 6
US-09-479-309-7
: Sequence 7, Application US/09479309
: Patent No. 6110691
: GENERAL INFORMATION:
: APPLICANT: Wang, Xiaodong
: APPLICANT: Du, Chunying
: TITLE OF INVENTION: Activators of Caspases
: FILE REFERENCE: UTS0630
: CURRENT APPLICATION NUMBER: US/09/479,309
: CURRENT FILING DATE: 2000-01-06
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 7
: LENGTH: 720
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-7
```

Query Match 84.8%; Score 71.2; DB 3; Length 720;
Best Local Similarity 90.5%; Pred. No. 1.9e-17;
Matches 76; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ACTTCATTTCTTCAGTACAGACAGTGTTCCTGTCCTGTTGCGCTAATTAGAAG 60
|||||
Db 37 ACTTCATTTCTTCAGTACAGACAGTGTTCCTGTCCTGTTGCGCTAATTAGAAG 96
|||||

OY 61 CGGTGTTCTCAGCAATGATAGA 84
|||||
Db 97 CGGTGTTCTCAGCAATGATAGA 120
|||||

RESULT 7

US-09-479-309-8
; Sequence 8, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Many, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTS0650
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 8
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-479-309-8

Query Match 84.8%; Score 71.2; DB 3; Length 720;
Best Local Similarity 90.5%; Pred. No. 1.9e-17;
Matches 76; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ACTTCATTTCTTCAGTACAGACAGTGTTCCTGTCCTGTTGCGCTAATTAGAAG 60
|||||
Db 37 ACTTCATTTCTTCAGTACAGACAGTGTTCCTGTCCTGTTGCGCTAATTAGAAG 96
|||||

OY 61 CGGTGTTCTCAGCAATGATAGA 84
|||||
Db 97 CGGTGTTCTCAGCAATGATAGA 120
|||||

RESULT 8

US-08-937-102-1/c
; Sequence 1, Application US/08937102A
; Patent No. 5965134
; GENERAL INFORMATION:
; APPLICANT: Thiel, Heinz-Jurgen
; APPLICANT: Elpers, Knut
; APPLICANT: Pauly, Thomas
; TITLE OF INVENTION: T Cell Stimulating Protein of Pestivirus
; FILE REFERENCE: 1/94108
; CURRENT APPLICATION NUMBER: US/08/937,102A
; CURRENT FILING DATE: 1997-09-24
; EARLIER APPLICATION NUMBER: 08/693,247
; EARLIER FILING DATE: 1996-08-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3639
; TYPE: DNA
; ORGANISM: Classical Swine Fever Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3639)
US-08-937-102-1

Query Match 29.3%; Score 24.6; DB 2; Length 3639;
Best Local Similarity 70.2%; Pred. No. 6.9;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 17 ACAGACAGTGTTCCTGTCCTGTTGCGCTAATTAGAAGCGG 63
|||||
Db 1066 ACATACAGTGTTCCTGTCCTGTTGCGCTAATTAGAAGCGG 1020
|||||

RESULT 9

US-08-876-991-1/c
; Sequence 1, Application US/08876991
; Patent No. 5925360
; GENERAL INFORMATION:
; APPLICANT: Gregor Meyers, Tillmann R menapf,
; APPLICANT: Heinz-J rgen Thiel
; TITLE OF INVENTION: Hog cholera virus vaccine and diagnostic
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Organon Teknika Corporation
; ADDRESSEE: Biotechnology Research Institute
; STREET: 1330-A Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/876,991
; FILING DATE: 16-JUN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US/08/747,577
; FILING DATE:
; APPLICATION NUMBER: US/08/650,584
; FILING DATE:
; APPLICATION NUMBER: US/08/469,702
; FILING DATE:
; APPLICATION NUMBER: US/08/123,596
; FILING DATE:
; APPLICATION NUMBER: 07/797,554
; FILING DATE: 22-NOV-1991
; APPLICATION NUMBER: US 07/494,991
; FILING DATE: 16-MAR-1990
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: William M. Blackstone
; REGISTRATION NUMBER: 29,772
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12284 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Hog cholera virus
; STRAIN: Alfort
; CELL LINE: PK 15 and 38A1D
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 364..12060
; OTHER INFORMATION: /label= 435_kDA_protein
; FEATURE:
; NAME/KEY: primer_bind

| | | | | | |
|--|--|-------------------------|---------------|----------------------------------|----------------------------------|
| | CELL LINE: | PK 15 | and | 38A1D | |
| | FEATURE: | | | | |
| | NAME/KEY: | CDS | | | |
| | LOCATION: | 364..12060 | | | |
| | OTHER INFORMATION: | /label= 435_kDA_protein | | | |
| | FEATURE: | | | | |
| | NAME/KEY: | primer_bind | | | |
| | LOCATION: | complement (2587..2619) | | | |
| | OTHER INFORMATION: | /label= primer_1 | | | |
| | FEATURE: | | | | |
| | NAME/KEY: | primer_bind | | | |
| | LOCATION: | complement (2842..2880) | | | |
| | OTHER INFORMATION: | /label= primer_2 | | | |
| | FEATURE: | | | | |
| | NAME/KEY: | variation | | | |
| | LOCATION: | replace(127, "c") | | | |
| | FEATURE: | | | | |
| | NAME/KEY: | variation | | | |
| | LOCATION: | replace(1522, "g") | | | |
| | FEATURE: | | | | |
| | NAME/KEY: | variation | | | |
| | LOCATION: | replace(10989, "t") | | | |
| | US-09-059-853-1 | | | | |
| | Query Match | 29.3%; | Score 24.6; | DB 2; | Length 12284; |
| | Best Local Similarity | 70.2%; | Pred. No. 10; | | |
| | Matches | 33; | Conservative | 0; | Mismatches 14; Indels 0; Gaps 0; |
| OY | 17 ACAACAGCTGTTGTGTCGTCCTGCTTGCGTACTTAAGAAGCGG 63 | | | | |
| | | | | | |
| Db | 4819 ACATACAGTGTGGTTCTTCTACTTAGTGTGCTGCCTTAAGCCG 4773 | | | | |
| | | | | | |
| RESULT 11 | | | | | |
| US-09-130-339-8 | | | | | |
| ; Sequence 8, Application US/09130339 | | | | | |
| ; Patent No. 6358701 | | | | | |
| ; GENERAL INFORMATION: | | | | | |
| ; APPLICANT: Cully, Doris F. | | | | | |
| ; APPLICANT: Pareess, Philip S. | | | | | |
| ; APPLICANT: Warmke, Jeffrey W. | | | | | |
| ; APPLICANT: Etter, Adrian | | | | | |
| ; APPLICANT: Cohen, Charles J. | | | | | |
| ; APPLICANT: Brochu, Richard M. | | | | | |
| ; TITLE OF INVENTION: DNA MOLECULES ENCODING CTENOCEPHALIDES | | | | | |
| ; FILE REFERENCE: 20029 | | | | | |
| ; CURRENT APPLICATION NUMBER: US/09/130,339 | | | | | |
| ; CURRENT FILING DATE: 1998-08-06 | | | | | |
| ; NUMBER OF SEQ ID NOS: 8 | | | | | |
| ; SOFTWARE: FastSeq for Windows Version 3.0 | | | | | |
| ; SEQ ID NO 8 | | | | | |
| ; LENGTH: 532 | | | | | |
| ; TYPE: DNA | | | | | |
| ; ORGANISM: ctenocephalides felis | | | | | |
| US-09-130-339-8 | | | | | |
| Query Match | 28.8%; | Score 24.2; | DB 4; | Length 532; | |
| Best Local Similarity | 66.0%; | Pred. No. 5.2; | | | |
| Matches | 35; | Conservative | 0; | Mismatches 18; Indels 0; Gaps 0; | |
| OY | 8 TCTTCAGGTACAGCAGCACTGTTGTGTCGTCGTCGTCGCTAACTTAAGAAG 60 | | | | |
| | | | | | |
| Db | 234 TCTTCAGGAGTGGACAGCATTAAGTTGCTCTACTTTGATGAACAAGCTTAAG 286 | | | | |
| | | | | | |
| RESULT 12 | | | | | |
| US-08-961-527-119/c | | | | | |
| ; Sequence 119, Application US/08961527 | | | | | |
| ; Patent No. 6420135 | | | | | |
| ; APPERAL INFORMATION: | | | | | |
| ; APPLICANT: Charles Kunsch | | | | | |
| ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences | | | | | |

```

Query Match          28.8%; Score 24.2; DB 4; Length 1968;
Best Local Similarity 59.4%; Pred. No. 8;
Matches 41; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

OY      16  TACAGACAGGTGTGTGTGTCCTGCTTGGCTTAACCTTGAAGACGCGTGTTCACGAA 75
          |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      1645  TACAGTCACACCTTGCTGCTCCCTGTATGTGCTGACTTGAATAATCGAGCGATCTGGAA 1586

OY      76  TTGATTAAGA 84
          |||  |  |
DB      1585  GATATCANA 1577

RESULT 13
US-09-039-198A-1
: Sequence 1, Application US/09039198A
: Patent No. 6200951
: GENERAL INFORMATION:
: APPLICANT: Gray, Patrick W.
: APPLICANT: Tjoelker, Larry W.
: TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker Drive/6300 Sears Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/039,198A
: FILING DATE:
: CLASSIFICATION: 435
:

```

```

OY      8 TCTTCAGGACAGACAGTGTGGTGCTGCCCTGTGTGCGCTAACCTTTAGA 58
        ||||| |||| | |||| | | | | ||||| |||| | |||| |
Db     975 TCTTCGCGGACACACAGTGGTGCGCTTTGATGATGCTGAGAGCTTCAAAA 1025

Query Match          28.3%; Score 23.8; DB 4; Length 1636;
Best Local Similarity    66.7%; Pred. No. 11;
Matches   34; Conservative    0; Mismatches   17; Indels    0; Gaps    0.

RESULT 14
US-08-877-599-1
: Sequence 1, Application US/08877599
: Patent No. 6372212
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,618
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rip-laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/33994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 16, 2003, 03:12:06 ; Search time 15.1045 Seconds
(without alignments)
622.464 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 327
Sequence: 1 acctcatctcttcagtgtaacg.....ccttagtgtagtaagaacatgca 184

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 280518

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09939293/runat_12022003.170356.8661/app_query.fasta.1.590
-DB=published.Applications_AA -QMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEARSTIE=500 -MINLEN=0
-MAXLEN=2000000000 -USBR=US09939293.qcgn_1.1.6.0/runat_12022003.170356.8661
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 322 | 98.5 | 227 | 10 | US-09-925-297-591 |
| 2 | 245 | 74.9 | 237 | 10 | US-09-798-116-2 |
| 3 | 245 | 74.9 | 237 | 10 | US-09-798-116-4 |
| 4 | 229 | 70.0 | 84 | 10 | US-09-798-116-9 |

| | | | | | | |
|----|------|------|------|----|---------------------|-------------------|
| 5 | 188 | 57.5 | 202 | 10 | US-09-798-116-7 | Sequence 7, Appl |
| 6 | 86 | 26.3 | 30 | 10 | US-09-939-293-7 | Sequence 7, Appl |
| 7 | 86 | 26.3 | 35 | 10 | US-09-939-293-11 | Sequence 11, Appl |
| 8 | 86 | 26.3 | 39 | 10 | US-09-939-293-8 | Sequence 8, Appl |
| 9 | 86 | 26.3 | 40 | 10 | US-09-939-293-2 | Sequence 2, Appl |
| 10 | 70 | 21.4 | 15 | 9 | US-10-068-569-8 | Sequence 8, Appl |
| 11 | 65.5 | 20.0 | 197 | 10 | US-09-205-658-235 | Sequence 235, App |
| 12 | 56.5 | 17.3 | 46 | 10 | US-09-864-761-34652 | Sequence 3452, A |
| 13 | 56.5 | 17.6 | 816 | 10 | US-09-815-242-11849 | Sequence 11849, A |
| 14 | 55 | 16.8 | 1798 | 10 | US-09-938-275-9 | Sequence 9, Appl |
| 15 | 54.5 | 16.7 | 746 | 9 | US-10-106-092-2 | Sequence 2, Appl |
| 16 | 53.5 | 16.4 | 133 | 10 | US-09-764-864-1301 | Sequence 1301, Ap |
| 17 | 53 | 16.2 | 13 | 10 | US-09-798-116-20 | Sequence 20, Appl |
| 18 | 53 | 16.2 | 13 | 10 | US-09-798-116-22 | Sequence 22, Appl |
| 19 | 53 | 16.2 | 177 | 10 | US-09-798-116-8 | Sequence 8, Appl |
| 20 | 53 | 16.2 | 243 | 9 | US-09-738-626-5950 | Sequence 5950, Ap |
| 21 | 53 | 16.2 | 1404 | 10 | US-09-944-849-8 | Sequence 8, Appl |
| 22 | 52.5 | 16.1 | 529 | 9 | US-09-712-363-175 | Sequence 175, App |
| 23 | 52.5 | 16.1 | 1172 | 9 | US-09-974-298-56 | Sequence 56, Appl |
| 24 | 52.5 | 16.1 | 1172 | 10 | US-09-919-172-16 | Sequence 16, Appl |
| 25 | 52 | 15.9 | 52 | 10 | US-09-925-299-892 | Sequence 892, App |
| 26 | 52 | 15.9 | 144 | 10 | US-09-925-301-1538 | Sequence 1538, Ap |
| 27 | 52 | 16.2 | 266 | 10 | US-09-929-493-19 | Sequence 19, Appl |
| 28 | 52 | 16.2 | 779 | 10 | US-09-749-601A-12 | Sequence 12, Appl |
| 29 | 52 | 15.9 | 929 | 9 | US-10-060-841-3 | Sequence 3, Appl |
| 30 | 52 | 15.9 | 1116 | 9 | US-09-977-577-10 | Sequence 10, Appl |
| 31 | 52 | 15.9 | 1149 | 9 | US-09-977-577-11 | Sequence 11, Appl |
| 32 | 52 | 15.9 | 1151 | 9 | US-09-977-577-13 | Sequence 13, Appl |
| 33 | 52 | 15.9 | 1156 | 9 | US-09-977-577-12 | Sequence 12, Appl |
| 34 | 51.5 | 15.7 | 508 | 10 | US-09-796-256A-2 | Sequence 2, Appl |
| 35 | 51.5 | 15.9 | 1352 | 10 | US-09-764-885-146 | Sequence 146, App |
| 36 | 51 | 15.6 | 172 | 10 | US-09-815-153-8 | Sequence 8, Appl |
| 37 | 51 | 15.6 | 197 | 9 | US-09-854-133-206 | Sequence 206, App |
| 38 | 51 | 15.6 | 197 | 10 | US-09-738-973-206 | Sequence 206, App |
| 39 | 51 | 15.6 | 211 | 9 | US-09-731-449-14 | Sequence 14, Appl |
| 40 | 51 | 15.6 | 211 | 9 | US-09-731-449-41 | Sequence 41, Appl |
| 41 | 51 | 15.6 | 211 | 9 | US-09-731-449-49 | Sequence 49, Appl |
| 42 | 51 | 15.6 | 223 | 9 | US-10-062-523-19 | Sequence 19, Appl |
| 43 | 51 | 15.6 | 223 | 10 | US-09-764-864-1343 | Sequence 1343, Ap |
| 44 | 51 | 15.6 | 263 | 9 | US-09-731-449-48 | Sequence 48, Appl |
| 45 | 51 | 15.6 | | | | |

ALIGNMENTS

RESULT 1
US-09-925-297-591
; Sequence 591, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 591
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-591
Alignment Scores: 1.25e-35 Length: 227


```
Db      31 ArgCysPheSerGluLeuIleArgProTyrPHisLysThrMetLeuThrGlyPheGlyVal 50
QY      121 ACCCTGTCGCGGTCTCTATTCGACAGAAATCAGACCCATTCCTTGTAGTACGACA 180
Db      51 ThrLeuCysAlaValProIleAlaGlnLysSerGluProHisSerLeuSerGlnAla 70
QY      181 TTG 183
Db      71 Leu 71

RESULT 5
US-09-798-116-7
; Sequence 7, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
; APPLICANT: Verhagen, Anne Marie
; APPLICANT: Ekert, Paul
; APPLICANT: Vaux, David
; TITLE OF INVENTION: NO. US20020110851A1el Polypeptides, Modulatory Agents Therefor an
; FILE REFERENCE: 10338-004US
; CURRENT APPLICATION NUMBER: US/09/798,116
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: AU PQ5995/00
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-116-7

Alignment Scores:
Pred. No.:      1,76e-17      Length:      202
Score:          188.00       Matches:      36
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    57.49%       Indels:      0
Db:            10           Gaps:        0

US-09-939-293-1_COPY_56_239 (1-184) x US-09-798-116-7 (1-202)
QY      76 TTGATNAGCATGTCACAAACTGTGACATTGCTTGAGTACCCCTGTGCGGTT 135
Db      1 LeuIleArgProTyrPHisLysThrValThrIleGlyPheGlyValThrLeuGlyAlaVal 20
QY      136 CCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGACATTG 183
Db      21 ProIleAlaGlnLysSerGluProHisSerLeuSerGlnAlaLeu 36

RESULT 6
US-09-939-293-7
; Sequence 7, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemrl, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-7

Alignment Scores:
Pred. No.:      0.00078      Length:      30
Score:          86.00       Matches:      18
Percent Similarity: 100.00%   Conservative: 0
```

```
Best Local Similarity: 100.00%   Mismatches: 0
Query Match:      26.30%       Indels:      0
Db:              10           Gaps:        0

US-09-939-293-1_COPY_56_239 (1-184) x US-09-939-293-7 (1-30)
QY      130 GCGGTTCTTATTCACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGACATTG 183
Db      1 AlaValProIleAlaGlnLysSerGluProHisSerLeuSerGlnAlaLeu 18

RESULT 7
US-09-939-293-11
; Sequence 11, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemrl, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-11

Alignment Scores:
Pred. No.:      0.000806     Length:      35
Score:          86.00       Matches:      18
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    26.30%       Indels:      0
Db:            10           Gaps:        0

US-09-939-293-1_COPY_56_239 (1-184) x US-09-939-293-11 (1-35)
QY      130 GCGGTTCTTATTCACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGACATTG 183
Db      1 AlaValProIleAlaGlnLysSerGluProHisSerLeuSerGlnAlaLeu 18

RESULT 8
US-09-939-293-8
; Sequence 8, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemrl, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-8

Alignment Scores:
Pred. No.:      0.000825     Length:      39
Score:          86.00       Matches:      18
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    26.30%       Indels:      0
Db:            10           Gaps:        0

US-09-939-293-1_COPY_56_239 (1-184) x US-09-939-293-8 (1-39)
QY      130 GCGGTTCTTATTCACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGACATTG 183
```

```
Db 1 AAlaValProIleAlaGlnLysSerGluProHisSerLeuSerGluAlaLeu 18
RESULT 9
US-09-939-293-2
; Sequence 2, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemrl, Emdad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-2

Alignment Scores:
Pred. No.: 0.000829 Length: 40
Score: 86.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.30% Indels: 0
DB: 10 Gaps: 0

US-09-939-293-1_COPY_56_239 (1-184) x US-09-939-293-2 (1-40)
Qy 130 GCGGTTCTATTGCAGAGAATCAGAGCCTATTCCTTAGTGAAGCATTTG 183
Db 1 AAlaValProIleAlaGlnLysSerGluProHisSerLeuSerGluAlaLeu 18
RESULT 10
US-10-068-569-8
; Sequence 8, Application US/10068569
; Patent No. US20020160975A1
; GENERAL INFORMATION:
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemrl, Teresa
; APPLICANT: Alnemrl, Emdad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; FILE REFERENCE: 480140.475
; CURRENT APPLICATION NUMBER: US/10/068,569
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-068-569-8

Alignment Scores:
Pred. No.: 0.0993 Length: 15
Score: 70.00 Matches: 14
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 93.33% Mismatches: 0
Query Match: 21.41% Indels: 0
DB: 9 Gaps: 0

US-09-939-293-1_COPY_56_239 (1-184) x US-10-068-569-8 (1-15)
Qy 130 GCGGTTCTATTGCAGAGAATCAGAGCCTATTCCTTAGTGAAGCATTTG 174
Db 1 AAlaValProIleAlaGlnLysSerGluProHisSerLeuSerGluAlaLeu 15
RESULT 11
US-09-205-658-235
```

```
; Sequence 235, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/205,658
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857,076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888,534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 235
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-658-235

Alignment Scores:
Pred. No.: 0.697 Length: 197
Score: 65.50 Matches: 15
Percent Similarity: 51.02% Conservative: 10
Best Local Similarity: 30.61% Mismatches: 13
Query Match: 20.03% Indels: 11
DB: 10 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x US-09-205-658-235 (1-197)
Qy 67 TTTCAGAAATGATAGACCATGGCAGCAAACTGGAGCATTTGC----- 111
Db 9 PheSerGluIleMetGlnProTyrPalaGlnThrValValGlyAlaArgLeuGly 28
Qy 112 -----TTTGAGTAACCTGTGTGCGGTTCTCATTTGCAGAGAAATCA 153
Db 29 GlyIleProValGlyValValAlaValGlnThrArgThrValGluLeuSerValProAla 48
Qy 154 GAGCCTATTCCTTAGTGAAGCA 180
Db 49 AspProAlaAsnLeuAspSerGluAla 57
RESULT 12
US-09-864-761-34652
; Sequence 34652, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```

APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11849
LENGTH: 816
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11849

Alignment Scores:
Pred. No.: 15 6 length: 816
Score: 56.50 Matches: 15
Percent Similarity: 47.83% Conservative: 7
Best Local Similarity: 32.61% Mismatches: 21
Query Match: 17.60% Indels: 3
DB: 10 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x US-09-815-242-11849 (1-816)
QY 158 GGCTCGATTTCTGTGCATATGAGAACCGACAGAGGTTACTCCAAAGCCATCGTCACA 99
|||||:||||| ||||| :||
Db 644 glyceroluphecysargpneglythrclanserthr-----glnleuglylle 660
QY 98 GTTTTGGCCATGGTCTTTATCAATTCTGAGAAACACCGCTTTTAAAGTTAGCCACACA 39
||| ||| ||| ||| ||| :|||
Db 661 AspleuclunhisAspleuclunhisAspleuclunhisAspleuclunhisAspleuclunhis 680
QY 38 GGACACACAAACACACGTCT 21
||| :|||
Db 681 glycySproArGsnCys 686

RESULT 14
US-09-938-275-9
: Sequence 9, Application US/09938275
: Patent No. US2002011309X1
: GENERAL INFORMATION:
: APPLICANT: Gerardo Castillo
: APPLICANT: Alan Snow
: TITLE OF INVENTION: Therapeutic and Diagnostic Applications
: FILE REFERENCE: PROTEO.P03
: CURRENT APPLICATION NUMBER: US/09/938,275
: CURRENT FILING DATE: 2001-08-16
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 1798
: TYPE: PRT
: ORGANISM: Homo Sapiens
: PUBLICACION INFORMATION:

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 02:25:59 ; Search time 132.896 Seconds

(without alignments)
1423.433 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84
Sequence: 1 actctctcttcagctacg.....gtttctcagaattgataaga 84

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 84 | 100.0 | 227 | 22 | AAF64278 | |
| 2 | 84 | 100.0 | 325 | 21 | AAC07201 | Novel human polynu |
| 3 | 84 | 100.0 | 720 | 21 | AAA94860 | Human caspase acti |
| 4 | 84 | 100.0 | 1358 | 24 | ABK15451 | DNA encoding inhib |
| 5 | 83 | 98.8 | 911 | 21 | AAC98904 | Human pancreatic c |
| 6 | 77.6 | 92.4 | 720 | 21 | AAA94247 | Human caspase acti |
| 7 | 77.6 | 92.4 | 720 | 21 | AAA94248 | Human caspase acti |
| 8 | 77.6 | 92.4 | 720 | 21 | AAA94249 | Human caspase acti |
| 9 | 71.2 | 84.8 | 720 | 21 | AAA94250 | Human caspase acti |

| | | | | | | |
|----|------|------|--------|----|----------|--------------------|
| 10 | 71.2 | 84.8 | 720 | 21 | AAA95001 | Human caspase acti |
| 11 | 71.2 | 84.8 | 720 | 21 | AAA95002 | Human caspase acti |
| 12 | 64.8 | 77.1 | 432 | 22 | AAF65005 | Novel human polynu |
| 13 | 54 | 64.3 | 376 | 22 | AAF66561 | Novel human polynu |
| 14 | 50.2 | 59.8 | 383 | 22 | AAF66548 | Novel human polynu |
| 15 | 27.8 | 33.1 | 6694 | 24 | ABL70351 | Chemically treated |
| 16 | 27.8 | 33.1 | 6694 | 24 | AA561312 | Human gene regulat |
| 17 | 27.6 | 32.9 | 21387 | 22 | AA542161 | Genomic sequence # |
| 18 | 27 | 32.1 | 2375 | 23 | ABL25042 | Drosophila melanog |
| 19 | 26.6 | 31.7 | 3373 | 23 | ABL25474 | Drosophila melanog |
| 20 | 26.6 | 31.7 | 6314 | 23 | ABL54314 | Chemically treated |
| 21 | 26.6 | 31.7 | 18173 | 23 | ABU05188 | Drosophila melanog |
| 22 | 26.2 | 31.2 | 537 | 23 | AAC94557 | Cat flea bindgut a |
| 23 | 26 | 31.0 | 421 | 23 | AA582701 | DNA encoding novel |
| 24 | 26 | 31.0 | 541 | 21 | AAF10783 | Fusarium venenatum |
| 25 | 26 | 31.0 | 6298 | 21 | ABL32435 | Human immune syste |
| 26 | 26 | 31.0 | 21354 | 22 | AA546784 | Tumour suppressor |
| 27 | 26 | 30.7 | 1960 | 23 | ABL30147 | Drosophila melanog |
| 28 | 25.8 | 30.7 | 3960 | 23 | ABL30146 | Drosophila melanog |
| 29 | 25.8 | 30.7 | 12586 | 22 | AAK67076 | Human immune/thema |
| 30 | 25.8 | 30.7 | 56390 | 23 | ABL22618 | Drosophila melanog |
| 31 | 25.8 | 30.5 | 8147 | 24 | ABU32429 | Human immune syste |
| 32 | 25.6 | 30.5 | 20933 | 24 | ABO67124 | Human angiogenesis |
| 33 | 25.6 | 30.5 | 30310 | 24 | AAU38253 | Genomic DNA from t |
| 34 | 25.6 | 30.5 | 113515 | 24 | ABU34175 | Human immune syste |
| 35 | 25.6 | 30.2 | 1140 | 24 | ABO35697 | Oligonucleotide fo |
| 36 | 25.4 | 30.2 | 1140 | 24 | ABO35697 | Oligonucleotide fo |
| 37 | 25.4 | 30.2 | 1390 | 24 | ABK37935 | Human immune syste |
| 38 | 25.4 | 30.2 | 13584 | 24 | ABL32614 | Human immune syste |
| 39 | 25.4 | 30.2 | 20778 | 22 | AAK79819 | Human immune/thema |
| 40 | 25.4 | 30.2 | 151826 | 21 | AAF22281 | BAC containing rep |
| 41 | 25.4 | 30.0 | 1223 | 24 | ABL80355 | Human ovarian can |
| 42 | 25.2 | 30.0 | 1695 | 22 | AA526188 | Human CDNA encodin |
| 43 | 25.2 | 30.0 | 2232 | 21 | AAA37532 | DNA encoding yeast |
| 44 | 25.2 | 30.0 | 2232 | 21 | AAA37532 | DNA encoding yeast |
| 45 | 25.2 | 30.0 | 2232 | 24 | ABA91860 | Yeast dual substra |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAF64278 | standard; cDNA; 227 BP. |
| ID | AAF64278: |
| XX | 09-APR-2001 (first entry) |
| XX | Novel human polynucleotide, SEQ ID NO: 34. |
| XX | Human: cytostatic; gene therapy: colon cancer; prostate cancer; |
| XX | breast cancer; lung cancer; cancer detection; ss. |
| XX | Homo sapiens. |
| XX | WO200102568-A2. |
| XX | 11-JAN-2001. |
| XX | 30-JUN-2000; 2000WO-US18374. |
| XX | 02-JUL-1999; 99US-0142310. |
| XX | 02-JUL-1999; 99US-0142311. |
| XX | (CHIR) CHIRON CORP. |
| XX | (HYSE-) HYSEQ INC. |
| PI | Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A; |
| PI | Reinhard R, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R; |
| PI | Ckenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D; |
| PI | Kita D, Garcia V, Jones LW, Strache-Crain B; |

DR WPI: 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 549; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
SQ Sequence 227 BP; 42 A; 60 C; 68 G; 57 T; 0 other;
Query Match 100.0%; Score 84; DB 22; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.4e-19;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTTCATCTTCAGGACAGACAGTGTGTGCTGCTGCTGCTAAGTAAAG 60
DB 131 ACTTCATCTTCAGGACAGACAGTGTGTGCTGCTGCTGCTAAGTAAAG 190
QY 61 CGGTGTTCTCAGAAATTGATAGA 84
DB 191 CGGTGTTCTCAGAAATTGATAGA 214
RESULT 2
AAC07201
ID AAC07201 standard; cDNA; 325 BP.
AC AAC07201;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 11276.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
OS
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 11276; 71pp + CD-ROM; English.

XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
SQ Sequence 325 BP; 71 A; 75 C; 86 G; 92 T; 1 other;
Query Match 100.0%; Score 84; DB 21; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.5e-19;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTTCATCTTCAGGACAGACAGTGTGTGCTGCTGCTGCTAAGTAAAG 60
DB 88 ACTTCATCTTCAGGACAGACAGTGTGTGCTGCTGCTGCTAAGTAAAG 147
QY 61 CGGTGTTCTCAGAAATTGATAGA 84
DB 148 CGGTGTTCTCAGAAATTGATAGA 171
RESULT 3
AAA94860
ID AAA94860 standard; cDNA; 720 BP.
AC AAA94860;
XX
DT 23-FEB-2001 (first entry)
XX
DE Human caspase activator Smac coding sequence.
XX
KW Human; caspase activator; Smac; apoptosis; cancer; autoimmune disease;
KW neurodegenerative disease; mitochondria; ss.
XX
OS Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 1..720 /tag- a
FT /product= "Smac"
XX
PN US6110691-A.
XX
PD 29-AUG-2000. ✓
XX
PF 06-JAN-2000; 2000US-0479309.
XX
PR 06-JAN-2000; 2000US-0479309.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
PI Wang X, Du C;
XX
DR WPI: 2000-586350/55.
DR P-PSDB; AAB26210.
XX
PT Novel caspase regulatory polypeptides useful for screening binding
PT agents specific for the polypeptides which are useful for diagnosis and
PT also for treating apoptosis associated diseases -
XX
PS Claim 5; column 21-22; 16pp; English.
CC The present sequence is the coding sequence of the human Smac protein. It

| | |
|-----------------------|---|
| CC | was isolated by purifying the protein and searching a Hela cell cDNA library for sequences which bound to probes based upon it. Smac is a mitochondrial protein which is released into the cytosol during apoptosis, and acts as a caspase-3 activator. The coding sequence and its protein can be used to modulate the expression and function of caspases and their activators, and also can be used as drug targets and regulators to promote or inhibit apoptosis in the treatment of cancer and autoimmune and neurodegenerative diseases. |
| XX | |
| SQ | Sequence 720 BP; 206 A; 147 C; 195 G; 172 T; 0 other; |
| Query Match | 100.0%; Score 84; DB 21; Length 720; |
| Best Local Similarity | 100.0%; Pred. No. 1.8e-19; |
| Matches | 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0 |
| YY | 1 ACTTCATCTTTCAGGTACAGACAGTTTGTTGTTCTTCCTGGTGTGCCTTAAGAAG 60 37 ACTTCATCTTTCAGGTACAGACAGTTTGTTGTTCTTCCTGGTGTGCCTTAAGAAG 96 |
| DY | 61 CGGTGTTTCTCAGAAATTGATAAGA 84 97 CGGTGTTTCTCAGAAATTGATAAGA 120 |
| DB | |
| RESULT 4 | |
| ABK15451 | |
| ID | ABK15451 standard; DNA; 1358 BP. |
| XX | |
| AC | ABK15451; |
| XX | |
| DT | 18-JUN-2002 (first entry) |
| DE | |
| XX | DNA encoding inhibitor of apoptosis (IAP) protein Smac. |
| KW | Human; Inhibitor of apoptosis; IAP; Smac; Apoptosis; BID; BIR1; BIR2; KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy; KM neoplastic cell; tumour; gene; ds. KW Homo sapiens. |
| OS | Homo sapiens. |
| XX | |
| FH | Key Location/Qualifiers |
| FT | CDS 20..739 /tag= a |
| FT | /product= "Smac protein" |
| FT | |
| PN | WO200216418-A2. |
| XX | |
| PD | 28-FEB-2002. |
| XX | |
| PF | 24-AUG-2001; 2001WO-US26492. |
| XX | |
| PR | 24-AUG-2000; 2000US-227735P. |
| XX | |
| PA | (UJJE-) UNIV JEFFERSON THOMAS. |
| PI | AInemri ES; |
| DR | WPI: 2002-304115/34. |
| XX | P-PSDB; AAU78447. |
| PT | Novel Smac peptides and polynucleotides encoding the peptides, useful for stimulating apoptosis in neoplastic or tumour cell which overexpresses inhibitor of caspase, and for identifying apoptosis modulating compounds - |
| PS | Claim 1; Page 73-74; 78pp; English. |
| CC | The invention relates to an isolated Smac peptide or polypeptide (I) and an isolated nucleic acid (II) encoding (I). Also described is a method of identifying a compound that inhibits apoptosis, comprising: (a) separately contacting several cell populations expressing a cytosolic Smac (a Smac isoform that begins with MKSDRF sequence, replacing the mitochondrial targeting sequence (residues 1-55 of (I)), |

| | |
|-----------------------|--|
| CC | and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting domain) which can be tested for apoptotic inhibiting activity; |
| CC | (b) incubating the cell populations with a direct stimulus of the cell death pathway; and (c) measuring the specific apoptotic activity of the cell populations, where inhibition of the specific apoptotic activity is indicative that the compound is an inhibitor of apoptosis. (I) and (II) are useful for inducing apoptosis in a cell. The Smac polypeptide and polynucleotide are useful for stimulating apoptosis in a neoplastic or tumour cell which overexpresses an inhibitor of caspase-3, caspase-7 or caspase-9. Preferably, the cell overexpresses at least a portion of IAP. |
| CC | (I) is useful for identifying an inhibitor or enhancer of a caspase-mediated apoptosis which involves contacting a cell transformed or transfected with a vector expressing (I) with a candidate inhibitor or candidate enhancer; and detecting cell viability, where an increase in cell viability indicates the presence of an inhibitor and a decrease in cell viability indicates the presence of an enhancer. Optionally, the method involves detecting the presence of large and small caspase subunits after contacting cell transformed with the vector expressing (I), with the candidate compound. A decrease in processing indicates the presence of an inhibitor and an increase in the processing indicates the presence of an enhancer. Preferably, the large and small subunits of caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for identifying a compound that inhibits Smac binding to Smac-binding molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3, or a full-length IAP). (II) is useful in gene therapy techniques. The present sequence represents the coding sequence of Smac protein. |
| SQ | Sequence 1358 BP; 341 A; 314 C; 358 G; 345 T; 0 other: |
| Query Match | 100.0%; Score 84; DB 24; Length 1358; |
| Best Local Similarity | 100.0%; Pred. No. 2.1e-19; |
| Matches | 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| OY | 1 ACTGATTCCTTAGGACAGACAGTGTGTGGTTCCGTGTGCGCTAACTTAAGAAG 60 |
| DB | 56 ACTGATTCCTTAGGACAGACAGTGTTGTGTGTTCTGTGTGCGCTAACTTAAGAAG 115 |
| OY | 61 CGGTGTTCTCAGAAATTGATAAGA 84 |
| DB | 116 CGGTGTTCTCAGAAATTGATAAGA 139 |
| RESULT 5 | |
| AAC98904 | ID AAC98904 standard; cDNA; 911 BP. |
| XX | AAC98904; |
| AC | 09-MAR-2001 (first entry) |
| DT | |
| XX | |
| DE | Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:132. |
| XX | |
| KW | Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection; diagnosis; identification; cytosolic; neuroprotective; nootropic; immunomodulatory; relaxant; contraceptive; gynaecological; antineoplastic; cardiant; gene therapy; chromosome mapping; linkage analysis; tissue identification; tissue typing; forensic; neural; immune system; muscular; reproductive; gastrointestinal; pulmonary; cardiovascular; renal; proliferative; ss. |
| KW | Homo sapiens. |
| OS | |
| PN | WO200055320-A1. |
| XX | |
| PD | 21-SEP-2000. |
| XX | |
| PF | 08-MAR-2000; 2000WO-US05989. |
| XX | |
| PR | 12-MAR-1999; 99US-0124270. |
| XX | |
| PA | (HUMA-) HUMAN GENOME SCI INC. |

PI Rosen CA, Ruben SM;
 XX
 DR WPI: 2000-579444/54.
 DR P-PSDB: AAB54139.
 XX
 PT New nucleic acid that is a pancreatic cancer antigen for preventing,
 PT treating, or ameliorating a medical condition, particular pancreatic
 PT cancer, or for use in assays for diagnosing a pathological condition -
 XX
 PS Claim 1, Page 591, 1379pp; English.
 XX
 CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,
 CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
 CC gynaecological, cardiac and antiinflammatory activities, and can be used
 CC in gene therapy. The polynucleotide and proteins can be used for
 CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in chromosome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect and target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 911 BP; 239 A; 202 C; 253 G; 211 T; 6 other;

Query Match 98.8%; Score 83; DB 21; Length 911;
 Best Local Similarity 100.0%; Pred. No. 4.2e-19;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTTCAATTTCTCAGTACAGACAGTGTGTCGTCCTGCTGCTAACTTAAGAAGC 61
 DB 64 CTTCAATTTCTCAGTACAGACAGTGTGTCGTCGCTGCTAACTTAAGAAGC 123
 QY 62 GGTGTTTCTCAGAAATGATAAGA 84
 DB 124 GGTGTTTCTCAGAAATGATAAGA 146

RESULT 6
 AAA94247
 ID AAA94247 standard; DNA; 720 BP.
 XX
 AC AAA94247;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 3.
 XX
 KW Human; caspase activator; Smac; hybridising sequence; apoptosis; cancer;
 KW autoimmune disease; neurodegenerative disease; mitochondria; ss.
 XX
 OS Homo sapiens.
 XX
 PN US6110691-A.
 XX
 PD 29-AUG-2000.
 XX
 PF 06-JAN-2000; 2000US-0479309.
 XX
 PR 06-JAN-2000; 2000US-0479309.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.

XX
 PI Wang X, Du C;
 XX
 DR WPI: 2000-586350/55.
 XX
 PT Novel caspase regulatory polypeptides useful for screening binding
 PT agents specific for the polypeptides which are useful for diagnosis and
 PT also for treating apoptosis associated diseases -
 XX
 PS Disclosure: column 23-26; 16pp; English.
 XX
 CC The present sequence is a sequence which was shown to hybridise to the
 CC coding sequence of the human Smac protein. The Smac coding sequence
 CC was isolated by purifying the protein and searching a HeLa cell cDNA
 CC library for sequences which bound to probes based upon it. Smac is a
 CC mitochondrial protein which is released into the cytosol during
 CC apoptosis, and acts as a caspase-3 activator. The coding sequence and
 CC protein can be used to modulate the expression and function of caspases
 CC and their activators, and also can be used as drug targets and regulators
 CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
 CC and neurodegenerative diseases.

Sequence 720 BP; 207 A; 145 C; 190 G; 178 T; 0 other;

Query Match 92.4%; Score 77.6; DB 21; Length 720;
 Best Local Similarity 95.2%; Pred. No. 3e-17;
 Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACTTATTTCTCAGTACAGACAGTGTGTCGTCCTGCTGCTAACTTAAGAAG 60
 DB 37 ACTTATTTCTCAGTACAGACAGTGTGTCGTCGCTAACTTAAGAAG 96
 QY 61 CGGTGTTTCTCAGAAATGATAAGA 84
 DB 97 CGGTGTTTCTCAGAAATGATAAGA 120

RESULT 7
 AAA94248
 ID AAA94248 standard; DNA; 720 BP.
 XX
 AC AAA94248;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 4.
 XX
 KW Human; caspase activator; Smac; hybridising sequence; apoptosis; cancer;
 KW autoimmune disease; neurodegenerative disease; mitochondria; ss.
 XX
 OS Homo sapiens.
 XX
 PN US6110691-A.
 XX
 PD 29-AUG-2000.
 XX
 PF 06-JAN-2000; 2000US-0479309.
 XX
 PR 06-JAN-2000; 2000US-0479309.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Wang X, Du C;
 XX
 DR WPI: 2000-586350/55.
 XX
 PT Novel caspase regulatory polypeptides useful for screening binding
 PT agents specific for the polypeptides which are useful for diagnosis and
 PT also for treating apoptosis associated diseases -
 XX
 PS Disclosure: column 25-26; 16pp; English.
 XX
 CC The present sequence is a sequence which was shown to hybridise to the

CC coding sequence of the human Smac protein. The Smac coding sequence
CC was isolated by purifying the protein and searching a HeLa cell cDNA
CC library for sequences which bound to probes based upon it. Smac is a
CC mitochondrial protein which is released into the cytosol during
CC apoptosis, and acts as a caspase-3 activator. The coding sequence and
CC protein can be used to modulate the expression and function of caspases
CC and their activators, and also can be used as drug targets and regulators
CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
CC and neurodegenerative diseases.

XX Sequence 720 BP; 202 A; 154 C; 192 G; 172 T; 0 other;

XX Query Match 92.4%; Score 77.6; DB 21; Length 720;

XX Best Local Similarity 95.2%; Pred. No. 3e-17; Mismatches 4; Indels 0; Gaps 0;

XX Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACTTCATTCTTCAGTACAGACAGTGTGTCCTGTTGCTGCTACTTTAAGAAG 60
DB 37 ACTTCATTCTTCAGTACAGACAGTGTGTCCTGTTGCTGCTACTTTAAGAAG 96
OY 61 CGGTGTTCTCAGAAATGATAAGA 84
DB 97 CGGTGTTCTCAGAAATGATAAGA 120

RESULT 8

AAA94249

XX AAA94249 standard; DNA: 720 BP.

XX AAA94249;

XX 23-FEB-2001 (first entry)

XX Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 5.

XX Human: caspase activator; Smac; hybridising sequence; apoptosis; cancer;
XX autoimmune disease; neurodegenerative disease; mitochondria; ss.

XX Homo sapiens.

XX US6110691-A.

XX 29-AUG-2000.

XX 06-JAN-2000; 2000US-0479309.

XX 06-JAN-2000; 2000US-0479309.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Wang X, Du C;

XX WPI: 2000-586350/55.

XX Novel caspase regulatory polypeptides useful for screening binding
XX agents specific for the polypeptides which are useful for diagnosis and
XX also for treating apoptosis associated diseases -
XX Disclosure: column 25-28; 16pp: English.

XX The present sequence is a sequence which was shown to hybridise to the
XX coding sequence of the human Smac protein. The Smac coding sequence
XX was isolated by purifying the protein and searching a HeLa cell cDNA
XX library for sequences which bound to probes based upon it. Smac is a
XX mitochondrial protein which is released into the cytosol during
XX apoptosis, and acts as a caspase-3 activator. The coding sequence and
XX protein can be used to modulate the expression and function of caspases
XX and their activators, and also can be used as drug targets and regulators
XX to promote or inhibit apoptosis in the treatment of cancer and autoimmune
XX and neurodegenerative diseases.

XX Sequence 720 BP; 205 A; 151 C; 188 G; 176 T; 0 other;

XX Query Match 92.4%; Score 77.6; DB 21; Length 720;

XX Best Local Similarity 95.2%; Pred. No. 3e-17; Mismatches 4; Indels 0; Gaps 0;

XX Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACTTCATTCTTCAGTACAGACAGTGTGTCCTGTTGCTGCTACTTTAAGAAG 60
DB 37 ACTTCATTCTTCAGTACAGACAGTGTGTCCTGTTGCTGCTACTTTAAGAAG 96
OY 61 CGGTGTTCTCAGAAATGATAAGA 84
DB 97 CGGTGTTCTCAGAAATGATAAGA 120

RESULT 9

AAA94250

XX AAA94250 standard; DNA: 720 BP.

XX AAA94250;

XX 23-FEB-2001 (first entry)

XX Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 6.

XX Human: caspase activator; Smac; hybridising sequence; apoptosis; cancer;
XX autoimmune disease; neurodegenerative disease; mitochondria; ss.

XX Homo sapiens.

XX US6110691-A.

XX 29-AUG-2000.

XX 06-JAN-2000; 2000US-0479309.

XX 06-JAN-2000; 2000US-0479309.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Wang X, Du C;

XX WPI: 2000-586350/55.

XX Novel caspase regulatory polypeptides useful for screening binding
XX agents specific for the polypeptides which are useful for diagnosis and
XX also for treating apoptosis associated diseases -
XX Disclosure: column 27-28; 16pp: English.

XX The present sequence is a sequence which was shown to hybridise to the
XX coding sequence of the human Smac protein. The Smac coding sequence
XX was isolated by purifying the protein and searching a HeLa cell cDNA
XX library for sequences which bound to probes based upon it. Smac is a
XX mitochondrial protein which is released into the cytosol during
XX apoptosis, and acts as a caspase-3 activator. The coding sequence and
XX protein can be used to modulate the expression and function of caspases
XX and their activators, and also can be used as drug targets and regulators
XX to promote or inhibit apoptosis in the treatment of cancer and autoimmune
XX and neurodegenerative diseases.

XX Sequence 720 BP; 208 A; 143 C; 193 G; 176 T; 0 other;

XX Query Match 84.8%; Score 71.2; DB 21; Length 720;

XX Best Local Similarity 90.5%; Pred. No. 4.9e-15; Mismatches 8; Indels 0; Gaps 0;

XX Matches 76; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ACTTCATTCTTCAGTACAGACAGTGTGTCCTGTTGCTGCTACTTTAAGAAG 60
DB 37 ACTTCATTCTTCAGTACAGACAGTGTGTCCTGTTGCTGCTACTTTAAGAAG 96
OY 61 CGGTGTTCTCAGAAATGATAAGA 84
DB 97 CGGTGTTCTCAGAAATGATAAGA 120

```
RESULT 10
AAA95001
ID AAA95001 standard; DNA: 720 BP.
XX
XX
AC AAA95001;
XX
XX 23-FEB-2001 (first entry)
XX
DE Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 7.
XX
XX Human; caspase activator; Smac; hybridising sequence; apoptosis; cancer;
XX
XX autoimmune disease; neurodegenerative disease; mitochondria; ss.
XX
OS Homo sapiens.
XX
XX US6110691-A.
XX
XX 29-AUG-2000.
XX
XX 06-JAN-2000; 2000US-0479309.
XX
XX 06-JAN-2000; 2000US-0479309.
XX
XX 06-JAN-2000; 2000US-0479309.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
XX Wang X, Du C;
XX
XX WPI; 2000-586350/55.
XX
XX Novel caspase regulatory polypeptides useful for screening binding
XX
XX agents specific for the polypeptides which are useful for diagnosis and
XX
XX also for treating apoptosis associated diseases -
XX
XX Disclosure; column 27-28; 16pp; English.
XX
XX The present sequence is a sequence which was shown to hybridise to the
XX
XX coding sequence of the human Smac protein. The Smac coding sequence
XX
XX was isolated by purifying the protein and searching a HeLa cell cDNA
XX
XX library for sequences which bound to probes based upon it. Smac is a
XX
XX mitochondrial protein which is released into the cytosol during
XX
XX apoptosis, and acts as a caspase-3 activator. The coding sequence and
XX
XX protein can be used to modulate the expression and function of caspases
XX
XX and their activators, and also can be used as drug targets and regulators
XX
XX to promote or inhibit apoptosis in the treatment of cancer and autoimmune
XX
XX and neurodegenerative diseases.
XX
XX Sequence 720 BP; 199 A; 162 C; 191 G; 168 T; 0 other;
XX
XX Query Match 84.8%; Score 71.2; DB 21; Length 720;
XX
XX Best Local Similarity 90.5%; Pred. No. 4.9e-15;
XX
XX Matches 76; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Oy 1 ACTTATTTCTCAGGTACAGACAGTGTGTCCTGTTGGCTAAGCTTAAGAAG 60
Db 37 ACTTATTTCTTCCGTACACAGACAGTGTGTCCTGTTGGCTAAGCTTAAGAAG 96
Oy 61 CGGTGTTCTCAGAAATGATAAGA 84
Db 97 CGGTGTTCTCAGAAATGATAAGA 120
RESULT 11
AAA95002
ID AAA95002 standard; DNA: 720 BP.
XX
XX AAA95002;
XX
XX 23-FEB-2001 (first entry)
XX
XX Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 8.
XX
XX Human; caspase activator; Smac; hybridising sequence; apoptosis; cancer;
```

```
KW autoimmune disease; neurodegenerative disease; mitochondria; ss.
XX
XX Homo sapiens.
XX
XX US6110691-A.
XX
XX 29-AUG-2000.
XX
XX 06-JAN-2000; 2000US-0479309.
XX
XX 06-JAN-2000; 2000US-0479309.
XX
XX 06-JAN-2000; 2000US-0479309.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
XX Wang X, Du C;
XX
XX WPI; 2000-586350/55.
XX
XX Novel caspase regulatory polypeptides useful for screening binding
XX
XX agents specific for the polypeptides which are useful for diagnosis and
XX
XX also for treating apoptosis associated diseases -
XX
XX Disclosure; column 29-30; 16pp; English.
XX
XX The present sequence is a sequence which was shown to hybridise to the
XX
XX coding sequence of the human Smac protein. The Smac coding sequence
XX
XX was isolated by purifying the protein and searching a HeLa cell cDNA
XX
XX library for sequences which bound to probes based upon it. Smac is a
XX
XX mitochondrial protein which is released into the cytosol during
XX
XX apoptosis, and acts as a caspase-3 activator. The coding sequence and
XX
XX protein can be used to modulate the expression and function of caspases
XX
XX and their activators, and also can be used as drug targets and regulators
XX
XX to promote or inhibit apoptosis in the treatment of cancer and autoimmune
XX
XX and neurodegenerative diseases.
XX
XX Sequence 720 BP; 218 A; 150 C; 184 G; 168 T; 0 other;
XX
XX Query Match 84.8%; Score 71.2; DB 21; Length 720;
XX
XX Best Local Similarity 90.5%; Pred. No. 4.9e-15;
XX
XX Matches 76; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Oy 1 ACTTATTTCTCAGGTACAGACAGTGTGTCCTGTTGGCTAAGCTTAAGAAG 60
Db 37 ACTTATTTCTCAGGTACAGACAGTGTGTCCTGTTGGCTAAGCTTAAGAAG 96
Oy 61 CGGTGTTCTCAGAAATGATAAGA 84
Db 97 CGGTGTTCTCAGAAATGATAAGA 120
RESULT 12
AAF65005
ID AAF65005 standard; cDNA: 432 BP.
XX
XX AAF65005;
XX
XX 09-APR-2001 (first entry)
XX
XX Novel human polynucleotide, SEQ ID NO: 761.
XX
XX Human; cytosolic; gene therapy; colon cancer; prostate cancer;
XX
XX breast cancer; lung cancer; cancer detection; ss.
XX
XX Homo sapiens.
XX
XX WO200102568-A2.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18374.
XX
XX 02-JUL-1999; 99US-0142310.
XX
XX 02-JUL-1999; 99US-0142311.
```

XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Rita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 654; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 432 BP; 118 A; 84 C; 104 G; 126 T; 0 other;
XX
Query Match 77.1%; Score 64.8; DB 22; Length 432;
Best Local Similarity 85.7%; Pred. No. 7.3e-13;
Matches 72; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
XX
QY 1 ACTTCATTCCTTCAGGTACAGACAGTGTGTCGTCGCTTAAGAG 60
DB 57 ACTTCATTCCTTCAGGTACAGACAGTGTGTCGCTTAAGAG 116
QY 61 CGGTGTTCTCAGATTTGATTAAGA 84
DB 117 CGTGTGTTCTCAGATTTGATTAAGA 140
XX
RESULT 13
AAF6561
ID AAF6561 standard; cDNA; 376 BP.
XX
AC AAF6561;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 2317.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000MO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Rita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 883; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 376 BP; 105 A; 70 C; 92 G; 108 T; 1 other;
XX
Query Match 64.3%; Score 54; DB 22; Length 376;
Best Local Similarity 98.2%; Pred. No. 3.9e-09;
Matches 54; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 30 GTGTGTTCCGTGTGCTTAAGAGCGGTGTTCTCAGATTTGATTAAGA 84
DB 9 GTGTGTTCCGTGTGCTTAAGAGCGGTGTTCTCAGATTTGATTAAGA 63
XX
RESULT 14
AAF6548
ID AAF6548 standard; cDNA; 383 BP.
XX
AC AAF6548;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 2304.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000MO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Rita D, Garcia V, Jones LW, Strache-Crain B;

XX WPI: 2001-091805/10.
 XX Library of polynucleotides for diagnosing a cancerous state of a
 PT mammalian cell and detecting cancer, particularly of the colon or
 PT prostate, comprises 3351 human polynucleotide sequences -
 XX
 PS Claim 9, Page 881-882; 1046pp; English.
 XX
 CC The present sequence is one of 3351 sequences in a library of human
 CC polynucleotides. The library is used to detect differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell and can
 CC detect colon, prostate, breast and lung cancer. The library can be used
 CC to produce probes for detection of mRNA and to produce additional copies
 CC of the polynucleotides. The probes can be used for chromosome mapping of
 CC the polynucleotide and for detection of transcription levels. Ribozymes
 CC or antisense oligonucleotides can be generated. The polynucleotides and
 CC their gene products are used as genetic or biochemical markers (e.g. in
 CC blood or tissues) that will detect the earliest changes along the
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and
 CC preventive interventions. The polynucleotides, polypeptides and
 CC antibodies against them can be used in pharmaceutical compositions to
 CC treat the cancers and proliferative disorders such as neoplasia,
 CC dysplasia and hyperplasia.
 CC
 SQ Sequence 383 BP; 108 A; 72 C; 90 G; 113 T; 0 other;
 Query Match 59.8%; Score 50.2; DB 22; Length 383;
 Best Local Similarity 94.5%; Pred. No. 8.1e-08;
 Matches 52; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 30 GGTGTTCTCTGTTGGCTAACTTTAGAGACGGCTTTCTCTCAGATTGATAGA 84
 DB 9 GGTGTTCTCTGTTGGCTAACTTTAGAGATCGTGTCTCTCAGATTGATAGA 63
 RESULT 15
 ID ABL70351
 AC ABL70351 standard; DNA; 6694 BP.
 XX
 AC ABL70351;
 DT 01-JUL-2002 (first entry)
 XX
 DE Chemically treated cell signalling DNA sequence#121.
 XX
 KW Cell signalling; cytosine methylation; cell signalling disease;
 KW cancer; tumour; cytostatic; ds.
 XX
 OS Unidentified.
 XX
 PN WO200202807-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-EP07471.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Olek A, Plepenbrock C, Berlin K;
 XX
 DR WPI: 2002-154758/20.
 XX
 XX Nucleic acid, useful for diagnosis and therapy of diseases associated
 PT with cell signalling e.g. cancer, comprises chemically modified genomic
 PT sequences of genes associated with cell signalling -
 XX
 PS Claim 1; SEQ ID NO 241; 24pp+sequence listing; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of at least

CC 18 bases of a segment of chemically pretreated DNA of genes associated
 CC with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytostatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or RNA-oligomers for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-ABL70626 represent chemically pre-treated
 CC genomic DNA's of genes associated with cell signalling.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 SQ Sequence 6694 BP; 1520 A; 121 C; 1737 G; 3316 T; 0 other;
 Query Match 33.1%; Score 27.8; DB 24; Length 6694;
 Best Local Similarity 59.5%; Pred. No. 9.2;
 Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 QY 3 TTCATTCTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAACTTTAGAGACG 62
 DB 3485 TTCGTTTTCAGGTTATATATTTTGTGTTTACGTTTTCAGTATGGAGATATAG 3544
 QY 63 GTGTTCTCAGATTGATA 81
 DB 3545 GTGTTGTTATATATATGTA 3563

Search completed: February 15, 2003, 12:12:24
 Job time : 136.896 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:10:33 : Search time 34 Seconds

(without alignments)
15.677 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues 11315

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.GeneSeq_101002.*

- 1: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 19 | 100.0 | 4 | 23 | ABB76220 |
| 2 | 19 | 100.0 | 4 | 23 | AAU97413 |
| 3 | 19 | 100.0 | 4 | 23 | AAU78441 |
| 4 | 19 | 100.0 | 4 | 23 | AAU78484 |
| 5 | 15 | 78.9 | 4 | 18 | AAW26094 |
| 6 | 15 | 78.9 | 4 | 19 | AAW51454 |
| 7 | 15 | 78.9 | 4 | 19 | AAW52389 |
| 8 | 15 | 78.9 | 4 | 20 | AAV39841 |
| 9 | 15 | 78.9 | 4 | 20 | AAV05586 |
| 10 | 15 | 78.9 | 4 | 21 | AAW28969 |

| | | | | | | |
|----|----|------|---|----|----------|--------------------|
| 11 | 15 | 78.9 | 4 | 21 | AAW83422 | Elastin derived pe |
| 12 | 15 | 78.9 | 4 | 21 | AAW87713 | Feline human TRFP |
| 13 | 15 | 78.9 | 4 | 21 | AAV90139 | TRFP derived pepti |
| 14 | 15 | 78.9 | 4 | 21 | AAV51510 | Human TRFP derived |
| 15 | 15 | 78.9 | 4 | 23 | AAU97418 | Synthetic tetrapep |
| 16 | 15 | 78.9 | 4 | 23 | AAU78444 | Inhibitor of apopt |
| 17 | 15 | 78.9 | 4 | 23 | AAU78445 | Inhibitor of apopt |
| 18 | 15 | 78.9 | 4 | 23 | AAU78495 | His-4 AV peptide |
| 19 | 15 | 78.9 | 4 | 23 | AAE16681 | Ligamentum nuchae |
| 20 | 14 | 73.7 | 3 | 13 | AAE29095 | Elastase-inhibitor |
| 21 | 14 | 73.7 | 3 | 14 | AAE38409 | Elastase-inhibitor |
| 22 | 14 | 73.7 | 4 | 5 | AAE40002 | Lipopeptide protea |
| 23 | 14 | 73.7 | 4 | 5 | AAE40457 | N-terminal polyhyd |
| 24 | 14 | 73.7 | 4 | 12 | AAE10886 | Peptide component |
| 25 | 14 | 73.7 | 4 | 12 | AAE11222 | Ketone analogue pr |
| 26 | 14 | 73.7 | 4 | 13 | AAE29963 | Elastase-inhibiti |
| 27 | 14 | 73.7 | 4 | 14 | AAE32269 | Morpholino urea de |
| 28 | 14 | 73.7 | 4 | 14 | AAE38407 | Elastase inhibitor |
| 29 | 14 | 73.7 | 4 | 15 | AAE46224 | Serine protease in |
| 30 | 14 | 73.7 | 4 | 15 | AAE53780 | Sequence of intern |
| 31 | 14 | 73.7 | 4 | 15 | AAE62048 | 21-Valinate ester |
| 32 | 14 | 73.7 | 4 | 16 | AAE77199 | Cell proliferation |
| 33 | 14 | 73.7 | 4 | 17 | AAE85708 | Degradable peptide |
| 34 | 14 | 73.7 | 4 | 18 | AAW52604 | Serine protease-in |
| 35 | 14 | 73.7 | 4 | 18 | AAW52607 | pp elastase inhibi |
| 36 | 14 | 73.7 | 4 | 18 | AAW29408 | Neutrophil elastas |
| 37 | 14 | 73.7 | 4 | 18 | AAW26090 | Substrate #13 for |
| 38 | 14 | 73.7 | 4 | 19 | AAW51759 | Indicator used to |
| 39 | 14 | 73.7 | 4 | 19 | AAW51447 | Indicator for dete |
| 40 | 14 | 73.7 | 4 | 19 | AAW51609 | Peptide conjugated |
| 41 | 14 | 73.7 | 4 | 19 | AAW52433 | Loop region used i |
| 42 | 14 | 73.7 | 4 | 19 | AAW52384 | Beta-turn region u |
| 43 | 14 | 73.7 | 4 | 19 | AAW52371 | Beta-turn region u |
| 44 | 14 | 73.7 | 4 | 19 | AAW26954 | Elastase-inhibiti |
| 45 | 14 | 73.7 | 4 | 19 | | |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| ABB76220 | |
| ID | ABB76220 standard; Peptide; 4 AA. |
| AC | ABB76220; |
| DT | 09-AUG-2002 (first entry) |
| XX | |
| DE | Human smac (DIABLO) derived peptide. |
| XX | |
| KW | DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis; |
| KW | human; cancer; cytosolic; mutant; mutleln. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FT | Key |
| FT | Modified-site 4 Location/Qualifiers |
| XX | W0200230959-A2. |
| PN | |
| PD | 18-APR-2002. |
| XX | |
| PF | 12-OCT-2001; 2001WO-US32121. |
| XX | |
| PR | 13-OCT-2000; 2000US-0687549. |
| XX | |
| PA | (ABBO) ABBOTT LAB. |
| XX | |
| PI | Fesik SW, Meadows RP, Betz SP, Liu Z, Olejniczak ET, Sun C; |
| XX | |
| DR | WPI; 2002-444169/47. |
| XX | |

PF Novel peptide derived from wild-type human second mitochondria derived
 PI activator of caspase protein useful for identifying candidate
 PI substances to kill cancerous cells -
 XX
 PS Example 1; Page 15; 26pp; English.
 XX
 CC The present sequence is a peptide derived from wild-type human
 CC second mitochondria derived activator of caspase (smac), also known
 CC as direct inhibitor of apoptosis binding protein with low PI
 CC (DIABLO). Claimed smac (DIABLO)-derived peptides (see ABN/6208-19)
 CC bind to the Bir2 and Bir3 domain of XIAP, an inhibitor of apoptosis
 CC protein (IAP) family member. They can be used to identify
 CC candidate substances which induce or promote apoptosis in cells.
 CC The assay involves determination of the ability of candidate
 CC compounds to disrupt the binding interaction between a smac
 CC (DIABLO) peptide and an IAP family member.
 CC
 SQ Sequence 4 AA;
 Query Match 100.0%; Score 19; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AVPI 4
 DB 1 AVPI 4
 IIII
 RESULT 2
 AAU97413
 ID AAU97413 standard; Peptide; 4 AA.
 XX
 AC AAU97413;
 XX
 DT 30-JUL-2002 (first entry)
 XX
 DE Human Smac (second mitochondria-derived activator of caspases) peptide.
 XX
 KW Human; synthetic tetrapeptide; mimetic; inhibitor of apoptosis; IAP;
 KW baculoviral IAP repeat; Bir; cell proliferative disease; cancer;
 KW Alzheimer's disease; second mitochondria-derived activator of caspases;
 KW stroke; arthritis; Smac.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 PN WO200226775-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 28-SEP-2001; 2001WO-US30567.
 XX
 PR 29-SEP-2000; 2000US-236574P.
 PR 20-DEC-2000; 2000US-256830P.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 PI Shi Y;
 XX
 DR WPI: 2002-416473/44.
 XX
 XX Novel synthetic tetrapeptide, its mimetic or a compound useful for
 PI rational drug design and in stimulating apoptosis in cell, binds an
 PI inhibitor of apoptosis protein and relieves inhibition of caspase -
 XX
 PS Claim 6; Page 43; 62pp; English.
 XX
 CC The present invention relates to a new synthetic tetrapeptide, its
 CC non-peptide or partial peptide mimetic or a compound that binds an
 CC inhibitor of apoptosis protein (IAP) and relieves IAP-mediated inhibition
 CC of caspase activity, where the tetrapeptide binds a surface groove within
 CC a baculoviral IAP repeat (Bir) domain of the IAP. The invention can be
 CC used for stimulating apoptosis in a cultured cell and for screening for a

CC compound that binds an IAP at a surface groove within a Bir domain.
 CC The synthetic tetrapeptide is a partial peptide or non-peptide mimetic
 CC and is useful in making a drug suitable for treating cell proliferative
 CC disease especially cancer in a mammal by promoting apoptosis in
 CC proliferatively diseased cells. Agonists identified by the invention are
 CC useful in treating cell proliferative diseases like cancer and
 CC antagonists identified by the invention are useful in treating
 CC Alzheimer's disease, stroke and arthritis. The present amino acid
 CC sequence represents the human Smac (second mitochondria-derived
 CC activator of caspases) peptide. This sequence is a synthetic
 CC tetrapeptide of the invention.
 CC
 SQ Sequence 4 AA;
 Query Match 100.0%; Score 19; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AVPI 4
 DB 1 AVPI 4
 IIII
 RESULT 3
 AAU78441
 ID AAU78441 standard; Peptide; 4 AA.
 XX
 AC AAU78441;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Inhibitor of apoptosis (IAP) protein Smac, peptide #1.
 XX
 KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; Bir1; Bir2;
 KW Bcl2 interacting domain; caspase; Bir domain; Bir3; gene therapy;
 KW neoplastic cell; tumour.
 XX
 OS Homo sapiens.
 OS
 PN WO200216418-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 24-AUG-2001; 2001WO-US26492.
 XX
 PR 24-AUG-2000; 2000US-227735P.
 XX
 PA (UYUE-) UNIV JEFFERSON THOMAS.
 XX
 PI Alnemri ES;
 XX
 DR WPI: 2002-304115/34.
 XX
 XX Novel Smac peptides and polynucleotides encoding the peptides, useful
 PI for stimulating apoptosis in neoplastic or tumour cell which
 PI overexpresses inhibitor of caspase, and for identifying apoptosis
 PI modulating compounds -
 XX
 PS Claim 8; Page 12; 78pp; English.
 XX
 CC The invention relates to an isolated Smac peptide or polypeptide (I)
 CC and an isolated nucleic acid (II) encoding (I). Also described is a
 CC method of identifying a compound that inhibits apoptosis, comprising:
 CC (a) separately contacting several cell populations expressing a
 CC cytosolic Smac (a Smac isoform that begins with MMSDRF sequence,
 CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
 CC domain) with a compound to be tested for apoptotic inhibiting activity;
 CC (b) incubating the cell populations with a direct stimulus of the cell
 CC death pathway; and (c) measuring the specific apoptotic activity of the
 CC cell populations, where inhibition of the specific apoptotic activity is
 CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
 CC are useful for inducing apoptosis in a cell. The Smac polypeptide and

CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
 CC tumour cell which overexpresses an inhibitor of caspase, where the
 CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
 CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
 CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
 CC mediated apoptosis which involves contacting a cell transformed or
 CC transfected with a vector expressing (I) with a candidate inhibitor or
 CC candidate enhancer; and detecting cell viability, where an increase in
 CC cell viability indicates the presence of an inhibitor and a decrease in
 CC cell viability indicates the presence of an enhancer. Optionally, the
 CC method involves detecting the presence of large and small caspase
 CC subunits after contacting cell transformed with the vector expressing
 CC (I), with the candidate compound. A decrease in processing indicates the
 CC presence of an inhibitor and an increase in the processing indicates the
 CC presence of an enhancer. Preferably, the large and small subunits of
 CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
 CC identifying a compound that inhibits Smac binding to Smac-binding
 CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
 CC or a full-length IAP). (II) is useful in gene therapy techniques. The
 CC present sequence represents the amino acid sequence of Smac peptide #1.
 CC
 XX
 SQ Sequence 4 AA;
 Query Match 100.0%; Score 19; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVPI 4
 DB 1 AVPI 4
 RESULT 4
 AAU78484
 ID AAU78484 standard; Peptide; 4 AA.
 AC AAU78484;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Smac-4 AV peptide.
 XX
 KW Apoptosis; cytosolic; apoptotic; AV peptide; melanoma; lymphoma;
 KW Inhibitor of Apoptosis protein; IAP; procaspase-3; tumour cell;
 KW breast cancer; prostate cancer; lung cancer; pancreatic cancer; smac-4;
 KW gastric cancer; colon cancer; ovarian cancer; renal cancer; hepatoma;
 KW sarcoma; smac; second mitochondria-derived activator of caspases.
 XX
 OS Synthetic.
 XX
 PN WO200216402-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 23-AUG-2001; 2001WO-US41869.
 XX
 PR 23-AUG-2000; 2000US-0645075.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Wang X, Du C;
 XX
 DR WPI: 2002-280909/32.
 XX
 PT Composition for enhancing the apoptosis of pathogenic cells,
 PT particularly tumour cells, e.g. breast cancer, prostate cancer, lung
 PT cancer, colon cancer, ovarian cancer or sarcoma, comprises apoptotic
 PT compounds -
 XX
 PS Example 9; Page 28; 40pp: English.
 CC This invention relates to a method for induction of apoptosis in
 CC pathogenic cells. The method comprises a novel pharmaceutical

CC composition which comprises an AV peptide in dosage form and a
 CC pharmaceutical carrier, where the AV peptide comprises a peptide that
 CC interacts with or inhibits the activity of an inhibitor of Apoptosis
 CC protein (IAP) as measured by IAP binding, procaspase-3 activation or
 CC promotion of apoptosis. The peptides of the invention are used to
 CC inhibit an inhibitor of Apoptosis protein (IAP). Compositions containing
 CC these peptides are useful for enhancing the apoptosis of pathogenic
 CC cells, particularly tumour cells, e.g. breast cancer, prostate cancer,
 CC lung cancer, pancreatic cancer, gastric cancer, colon cancer, ovarian
 CC cancer, renal cancer, hepatoma, melanoma, lymphoma or sarcoma. The
 CC composition is particularly useful for promoting cell death. The
 CC present sequence represents an AV peptide (smac-4) used to inhibit
 CC second mitochondria-derived activator of caspases (smac) using the
 CC method of the invention. Smac interacts with and eliminates the activity
 CC of a number of IAP's and as such inhibiting its activity allows the
 CC induction of apoptosis.
 CC
 XX
 SQ Sequence 4 AA;
 Query Match 100.0%; Score 19; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVPI 4
 DB 1 AVPI 4
 RESULT 5
 AAW26094
 ID AAW26094 standard; peptide; 4 AA.
 AC AAW26094;
 XX
 DT 05-NOV-1997 (first entry)
 XX
 DE Substrate #17 for serine protease.
 XX
 KW Serine protease; substrate; Streptomyces griseus; guanidine; pre-soak;
 KW cleaning composition; laundry detergent; additive composition; enzyme;
 KW dishwasher detergent; drain opener; urea; contact lens cleanser;
 KW proteinaceous stain.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 XX
 FT Modified-site 1 /note- "succinylated"
 FT FT /note- "para-nitroamylidated"
 FT Modified-site 4 /note- "para-nitroamylidated"
 XX
 FT US5646028-A.
 XX
 PN 08-JUL-1997.
 XX
 PD 18-JUN-1991; 91US-0718303.
 XX
 PF 18-JUN-1991; 91US-0718303.
 XX
 PR 06-NOV-1992; 92US-0973343.
 XX
 PR 18-AUG-1994; 94US-0292924.
 XX
 PR 17-OCT-1995; 95US-0544143.
 XX
 PA (CLR) CLOBOX CO.
 XX
 PI Leigh SD;
 XX
 DR WPI: 1997-362936/33.
 XX
 PT Serine protease from Streptomyces griseus ATCC 55178 - with good
 PT stability in presence of urea or guanidine, useful in cleaning
 PT compositions, including laundry and dishwashing detergents
 XX
 PS Example 1; Column 12; 16pp; English.

```

XX AAM26078-W26096 represent substrates for the serine protease of the
CC invention. The serine protease recognises these sequences, but is
CC specific for the sequence shown in AAM24567. The protease has the
CC N-terminal and C-terminal sequences represented by AAM24565 and AAM24566
CC respectively. The serine protease was isolated from Streptomyces griseus
CC variety alkaliphilus No. 33 (ATCC 55178). The protease has an apparent
CC molecular weight of 19 kD (by reducing sodium dodecylsulphate
CC polyacrylamide gel electrophoresis), and improved stability against urea
CC and guanidine. The protease is inhibited by phenylmethylsulphonyl
CC fluoride. The serine protease is useful in liquid or granular cleaning
CC compositions, specifically laundry detergents or additive compositions.
CC It is also useful in automatic dishwasher detergents, pre-soaks, drain
CC openers, contact lens cleansers etc. The protease has better activity
CC against proteinaceous stains than known enzymes and unusually high
CC stability in the presence of chaotropic agents.
XX
SQ Sequence 4 AA;
Query Match 78.9%; Score 15; DB 18; Length 4;
Best Local Similarity 75.0%; Pred. No. 7.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AVPI 4
   1 11
Db 1 AAP1 4
RESULT 6
AAM51454
ID AAM51454 standard; peptide; 4 AA.
XX
AC AAM51454;
XX
DT 02-SEP-1998 (first entry)
XX
DE Indicator for detection of leukocyte esterase activity in urine.
XX
KW Urine; indicator; detection; leukocyte esterase; activity;
XX white blood cell.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Methoxysuccinyl-suc-Ala"
FT Modified-site 4 /note= "Met-nitroanilide"
FT FT
XX
XX US5776780-A.
XX
XX 07-JUL-1998.
XX
XX 12-APR-1996; 96US-0631581.
XX
XX 12-APR-1996; 96US-0631581.
XX
XX 28-MAY-1993; 93US-0068956.
XX
XX 24-APR-1995; 95US-0429292.
XX
XX (CHIM-) CHIMERA RES & CHEM INC.
XX
XX Carter JM, Smith JV;
XX
XX WPI, 1998-398049/34.
XX
XX Automated measurement of white blood cell esterase activity in urine
XX i comprises adding reagent composition to the sample containing an
XX indicator, placing sample in a analyser and comparing absorbance
XX with a standard measurement
XX
XX Claim 11; Column 14; 8pp; English.
XX
XX The invention relates to a method for the measurement of white blood
CC

```

```

CC cell esterase activity in urine. It comprises: (1) placing an aliquot of
CC the urine in an automated analyser sampling cup; (1i) placing the cup in
CC a sampling tray with the automated analyser, transferring the urine to a
CC cuvette and injecting at least one reagent composition in an aqueous
CC medium into the cuvette, where the reagent composition comprises a
CC buffer to adjust the pH of the urine to 7-9, an activator and surfactant
CC selected from benzalkonium chloride, 2,5-dimethylbenzenesulphonic acid
CC and benzethonium chloride, and an indicator to determine leukocyte
CC esterase activity in the urine; (1ii) reading the aliquot of urine at
CC specified intervals, in accordance with a preprogrammed code introduced
CC into the automated analyser, at a preprogrammed monochromatically
CC specified wavelength, to compare absorbance of the urine sample and
CC reagent composition complex with that of a standard containing a known
CC concentration of leukocyte esterase and determining the quantitative
CC amount of leukocyte esterase in the patient's urine. The method provides
CC a fully automated method for determining white blood cells in urine. The
CC method is preferable to the old dipstick method as it reduces consumable
CC material and labour costs. It also offers increased accuracy, sensitivity
CC and reduction of interference by substances which affected prior art
CC tests. The present sequence represents a specifically claimed indicator.
XX
SQ Sequence 4 AA;
Query Match 78.9%; Score 15; DB 19; Length 4;
Best Local Similarity 50.0%; Pred. No. 7.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 AVPI 4
   1 11
Db 1 AIPW 4
RESULT 7
AAM52389
ID AAM52389 standard; peptide; 4 AA.
XX
AC AAM52389;
XX
DT 01-JUL-1998 (first entry)
XX
DE Beta-turn region used in cyclic peptide of the invention.
XX
XX Beta-turn region; cyclic peptide; antimicrobial; disinfectant; therapy;
XX preservation; amphipathic anti-parallel beta-sheet region; plant disease.
XX
XX Synthetic.
XX
XX WO9803192-A1.
XX
XX 29-JAN-1998.
XX
XX 23-JUL-1997; 97WO-US12974.
XX
XX 24-JUL-1996; 96US-0685589.
XX
XX (INTR-) INTRABIOFICS PHARM INC.
XX
XX Chang C, Chen J, Gu L;
XX
XX WPI, 1998-120472/11.
XX
XX New cyclic peptide(s) with antimicrobial activity - contain
XX amphipathic beta-sheet, loop and beta-turn regions, have better
XX activity, bioavailability and protease resistance than linear
XX analogues
XX
XX Claim 3; Page 149; 160pp; English.
XX
XX This sequence represents a beta-turn region used in a peptide of the
XX invention. The peptides are cyclic peptides (I), which have: (a) an
XX amphipathic anti-parallel beta-sheet region (SR), a loop region (LR) and
XX a beta-turn region (TR); (b) a net positive charge at physiological pH;
XX and (c) at least one basic amino acid (aa) in LR or TR. (I) are broad
CC

```

CC spectrum antimicrobials, specifically for use against *E. coli*,
CC *Pseudomonas aeruginosa*, methicillin-resistant *Staphylococcus aureus*
CC (MRSA), vancomycin-resistant *Enterococcus faecium* and
CC penicillin-resistant *Streptococcus pneumoniae*. More generally they are
CC active against Gram-positive or -negative bacteria, fungi, yeast and
CC protozoa. Apart from clinical uses, (I) are also used as disinfectants
CC and preservatives for medical equipment, foods, cosmetics etc., also for
CC treatment of plant diseases. Compared with non-cyclised analogues (i.e.
CC tachyplesin and protegrin type peptides), (I) and are more effective,
CC with better bioavailability and/or serum half-life (increased resistance
CC to proteolysis). They are more suitable for oral administration, can be
CC used at lower doses and are unlikely to induce development of resistant
CC strains.
XX
SQ Sequence 4 AA;

Query Match 78.9%; Score 15; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VPI 4
|||
Db 1 VPI 3

RESULT 8
AA39841
ID AA39841 standard; peptide: 4 AA.
XX
AC AA39841;
XX
DT 03-DEC-1999 (first entry)
XX
DE Elastin peptide fragment #15.
XX
KW Elastin peptide; mammalian tissue; hypertension; coronary heart disease;
KW arteriosclerosis; angina; coronary thrombosis; neovascularisation;
KW chronic obstructive pulmonary disease; restenosis post angioplasty;
KW tissue turgor; tissue elasticity; angiogenesis; therapy.
XX
OS Mammalia.
XX
PM WO945941-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US05496.
XX
PR 13-MAR-1998; 98US-0039308.
XX
PA (MRS-) MRS LLC.
XX
PI Sandberg LB, Roos PJ, Mltis TF;
XX
DR WPI, 1999-561616/47.
XX
PT Elastin peptide fragment compositions useful for treating hypertension,
XX coronary heart disease and arteriosclerosis -
XX
PS Claim 4; Page 21; 55pp; English.
XX
CC This sequence represents an elastin peptide fragment that is used in the
CC composition of the invention, which is useful in treating a condition of
CC mammalian tissue. The elastin peptide fragment composition is useful for
CC treating hypertension, coronary heart disease, arteriosclerosis, angina,
CC coronary thrombosis, chronic obstructive pulmonary disease and restenosis
CC post angioplasty. The mammalian tissue being treated is a mammalian blood
CC vessel. The composition is useful in improving tissue turgor. The
CC composition can be used in a method for enhancing tissue elasticity. The
CC method can be used to enhance tissue, especially skin, elasticity.
CC Increased elasticity and/or physical appearance of the skin gives a
CC smoother, healthier and youthful looking skin. The method also comprises
CC stimulating endogenous production of elastin. The method can be used to

CC enhance elasticity of blood vessels or lung tissue. Application of the
CC composition results in neovascularisation or angiogenesis.
XX
SQ Sequence 4 AA;

Query Match 78.9%; Score 15; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVP 3
|||
Db 1 AVP 3

RESULT 9
AA09586
ID AA09586 standard; peptide: 4 AA.
XX
AC AA09586;
XX
DT 21-JUL-1999 (first entry)
XX
DE Antimicrobial cyclic peptide beta-turn region SEQ ID NO:33.
XX
KW Cyclic peptide; antimicrobial; amphiphilic beta-sheet; biostatic;
KW biocidal; bacteria; yeast; Candida species; fungi; protozoa;
KW *Escherichia coli*; *Pseudomonas aeruginosa*; infection; preservative;
KW vancomycin-resistant *Enterococcus faecium*; disinfectant; food;
KW methicillin-resistant *Staphylococcus aureus*; medical equipment;
KW penicillin-resistant *Streptococcus pneumoniae*; cosmetic.
XX
OS Synthetic.
XX
PN WO921879-A1.
XX
PD 06-MAY-1999.
XX
PF 27-OCT-1997; 97WO-US19557.
XX
PR 27-OCT-1997; 97WO-US19557.
XX
PA (INTR-) INTRABIOPTICS PHARM INC.
XX
PI Chang C, Chen J, Gu L;
XX
DR WPI, 1999-312941/26.
XX
PT Antimicrobial cyclic peptides with amphiphilic beta-sheet region
XX
PS Claim 3; Page 150; 167pp; English.
XX
CC The present invention describes cyclic peptides (I): (1) comprising an
CC amphiphilic anti-parallel beta-sheet region (A), a loop region (B) and
CC a beta-turn (C); (ii) having net positive charge at physiological pH;
CC and (iii) including at least one basic amino acid (aa) in (B) or (C).
CC AA09554 to AA09683 and AA17301 to AA17385 represent specifically
CC claimed cyclic peptides and peptide segments. (I) are broad-spectrum
CC antimicrobials (biostatic or biocidal), effective against bacteria,
CC yeast (e.g. *Candida species*), fungi and protozoa. Particularly they
CC are used to control *Escherichia coli*, *Pseudomonas aeruginosa*,
CC vancomycin-resistant *Enterococcus faecium*, methicillin-resistant
CC *Staphylococcus aureus* and penicillin-resistant *Streptococcus*
CC *pneumoniae*. (A) are used to treat or prevent infections, in animals or
CC plants, also as preservatives and disinfectants for medical equipment,
CC foods, cosmetics, optionally as mixtures or in combination with other
CC antimicrobials. Compared with linear analogues, (I) are more effective
CC with better bioavailability and/or serum half-life (better resistance
CC to proteolysis, allowing lower doses and making them more suitable for
CC oral delivery). Since (I) are structurally related to naturally
CC occurring antimicrobial peptides, they are less likely to induce
CC development of resistant strains.
XX
SQ Sequence 4 AA;

Query Match 78.9%; Score 15; DB 20; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VPI 4
 |||
 Db 1 VPI 3

RESULT 10

AAB28969 standard; Peptide: 4 AA.

AC AAB28969;

DT 29-JAN-2001 (first entry)

DE Peptide encoded by oligonucleotide I.

KW Cat; allergy; human T cell reactive feline protein; hTRFP;

OS Felis sp.

OS Synthetic.

PN US6120769-A.

PD 19-SEP-2000.

PE 28-APR-1995; 95US-0431184.

PR 02-SEP-1994; 94US-0300928.

PR 03-NOV-1989; 89US-0431565.

PR 28-FEB-1991; 91US-0662276.

PR 13-DEC-1991; 91US-0807529.

PR 25-MAR-1992; 92US-0857311.

PR 15-MAY-1992; 92US-0884718.

PR 15-JAN-1993; 93US-0006116.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Geffer ML, Garman RD, Greenstein JL, Bond JF;

DR WPI: 2000-601477/57.

DR N-PSDB: AAC60134.

XX Detecting, preventing and treating sensitivity to cat protein allergen

PT comprises combining a biological sample with a human T cell reactive

PT feline protein and determining the extent of binding that occurs -

XX Example 25; Column 111-112; 106pp; English.

XX The present invention relates to the detection of sensitivity to a cat

CC protein allergen by combining a blood sample from a subject with a

CC peptide of human T cell reactive feline protein (hTRFP). This method

CC and the hTRFP peptides are useful for diagnosing, preventing and

CC treating cat allergies by reducing or abolishing an individual's

CC allergic response to a cat allergen. DNA encoding the TRFP may be

CC used as probes to locate equivalent sequences present in other species.

CC These may further be used to study the mechanism of immunotherapy of

CC cat allergy, and to design modified derivatives, analogues or

CC functional equivalents useful in immunotherapy. The present

CC sequence was used in the invention.

XX Sequence 4 AA;

SO Query Match 78.9%; Score 15; DB 21; Length 4;

Best Local Similarity 50.0%; Pred. No. 7.8e+05;

Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
 |||:

Db 1 ALPV 4

RESULT 11

AAV83422 standard; peptide: 4 AA.

AC AAV83422;

DT 29-AUG-2000 (first entry)

DE Elastin derived peptide.

KW Elastin; peptide; cosmetic; epidermis; skin; dermis;

KW dermatological; cutaneous; ageing; steroid; wrinkles; human.

OS Homo sapiens.

PN WO200028996-A1.

PD 25-MAY-2000.

PF 19-NOV-1999; 99WO-US273399.

PR 19-NOV-1998; 98US-0109246.

PA (MRSW-) MRS LLC.

PI Mlts TF, Sandberg LB, Roos PJ, Dray FJ;

DR WPI: 2000-399644/34.

PT Cosmetic or dermatological composition comprise 7-hydroxylated steroid

PT and optionally elastin peptide fragment, useful for improving tissue

PT texture

PS Disclosure; Page 42; 74pp; English.

XX A cosmetic or dermatological composition comprising a 7-hydroxylated

CC steroid and optionally an elastin peptide fragment can be used for

CC restoring and preventing cutaneous ageing in persons of a certain

CC age as well as providing protective effects in relation to UV,

CC firming up the skin and reducing wrinkles. Elastin derived peptides

CC are described in GENES50 records AAV83412-Y83447.

XX Sequence 4 AA;

SO Query Match 78.9%; Score 15; DB 21; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVP 3
 |||

Db 1 AVP 3

RESULT 12

AAV87713 standard; Protein: 4 AA.

AC AAV87713;

DT 22-AUG-2000 (first entry)

DE Feline human TRFP oligonucleotide I derived peptide.

KW T-cell reactive feline protein; TRFP; Fel d I; cat allergen; primer;

KW antiallergic; T cell stimulator; diagnostic; immunotherapy.

OS Felis sp.

PN US6048962-A.

PD 11-APR-2000.
 XX
 PF 27-APR-1995; 95US-0430014.
 XX
 PR 02-SEP-1994; 94US-0300928.
 PR 03-NOV-1989; 89US-0431565.
 PR 28-FEB-1991; 91US-0662276.
 PR 13-DEC-1991; 91US-0807529.
 PR 25-MAR-1992; 92US-0857311.
 PR 15-MAY-1992; 92US-0884718.
 PR 15-JAN-1993; 93US-0006116.
 PA
 XX (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Kuo M, Rogers BL, Gelfer ML, Morgenstern JP, Brauer AM,
 PI Greenstein JL, Griffith ID, Garman RD;
 DR WPI: 2000-316905/27.
 DR N-PSDB: AAM12272.
 XX
 PT New human T cell reactive feline protein useful for reducing or
 PT abolishing individual's allergic response to cat allergen comprising
 PT two different covalently linked peptide chains -
 XX
 PS Example 25; Column 111-112; 106pp; English.
 XX
 CC This invention describes a novel naturally occurring cat protein allergen
 CC (1), human T cell reactive feline protein (TRFP), comprising two
 CC different covalently linked peptide chains with a molecular weight of 20
 CC kD, 40 kD or 130 kD under non-reducing conditions and 5 kD or 10-18 kD
 CC under reducing conditions. The products of the invention have
 CC antiallergic activity and act as human T cell stimulators. TRFP is useful
 CC for reducing or preventing the adverse effects of cat allergens on cat
 CC allergic individuals and in ex vivo diagnostic tests to determine which
 CC peptides cause sensitivity so as to selectively use them to desensitize
 CC a cat sensitive individual. Purified TRFP is also useful for studying
 CC the mechanism of immunotherapy of cat allergy and to design modified
 CC derivatives, analogs or functional equivalents that are more useful in
 CC immunotherapy against cat allergy. DNA sequences encoding TRFP are
 CC useful as probes to locate equivalent sequences present in other species
 CC (goats, sheep, dogs, rabbits or horses) that may be useful in diagnostics
 CC and/or therapeutics. Fully defined and characterized TRFP provides
 CC complete and a very simple desensitization therapy. This sequence
 CC represents a peptide fragment of TRFP used in the design of a primer
 CC which is used in the construction and expression of peptides comprising
 CC two or more regions of TRFP in a method described in the invention.
 CC
 SQ Sequence 4 AA;
 XX
 Query Match 78.9%; Score 15; DB 21; Length 4;
 Best Local Similarity 50.0%; Pred. No. 7.8e+05;
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AVPI 4
 I:I:
 Db 1 ALPV 4
 XX
 RESULT 13
 AAY90139
 ID AAY90139 standard; Peptide; 4 AA.
 XX
 AC AAY90139;
 XX
 DT 13-JUL-2000 (first entry)
 XX
 DE TRFP derived peptide, peptide I.
 XX
 KM Cat; TRFP; human T-cell reactive feline protein; cat protein allergen;
 KM house dust; fel d I; cat allergy; Felis domesticus sensitivity; therapy;
 KM diagnosis; goat; sheep; horse; rabbit; dog.
 XX
 OS Felis domesticus.

XX
 PN US6025162-A.
 XX
 PD 15-FEB-2000.
 XX
 PF 28-APR-1995; 95US-0430944.
 XX
 PR 02-SEP-1994; 94US-0300928.
 PR 03-NOV-1989; 89US-0431565.
 PR 28-FEB-1991; 91US-0662276.
 PR 13-DEC-1991; 91US-0807529.
 PR 25-MAR-1992; 92US-0857311.
 PR 15-MAY-1992; 92US-0884718.
 PR 15-JAN-1993; 93US-0006116.
 PA
 XX (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Morgenstern JP, Griffith ID, Rogers BL;
 PI WPI: 2000-181812/16.
 DR
 XX
 PT New human T cell reactive feline protein, useful for desensitizing cat
 PT allergic individuals to cat allergens -
 XX
 PS Example 25; Fig 28; 108pp; English.
 XX
 CC This sequence is a peptide derived from the human T cell reactive feline
 CC protein (TRFP) of the invention, and was used to design a PCR
 CC primer. The protein is a cat protein allergen,
 CC and was isolated from a vacuum bag extract obtained by affinity
 CC purification of house dust collected from several homes with cats. TRFP
 CC is composed of two covalently linked peptide chains, and is also referred
 CC to as fel d I. TRFP and its peptides are useful for reducing or
 CC preventing the adverse effects that exposure to cat allergens normally
 CC has on cat allergic individuals (i.e. to desensitize individuals to cat
 CC allergens or block the effect of the allergens). TRFP is also used in
 CC methods of diagnosing sensitivity to Felis domesticus in an individual.
 CC DNA sequences encoding TRFP can be used as probes to locate equivalent
 CC sequences present in other species, e.g. goat, sheep, horse, rabbit and
 CC dog, that may be useful in diagnostic and/or therapeutic applications.
 CC
 SQ Sequence 4 AA;
 XX
 Query Match 78.9%; Score 15; DB 21; Length 4;
 Best Local Similarity 50.0%; Pred. No. 7.8e+05;
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AVPI 4
 I:I:
 Db 1 ALPV 4
 XX
 RESULT 14
 AAY51510
 ID AAY51510 standard; Protein; 4 AA.
 XX
 AC AAY51510;
 XX
 DT 22-MAY-2000 (first entry)
 XX
 DE Human TRFP derived peptide encoding oligonucleotide I.
 XX
 KM T-cell reactive feline protein; TRFP; T cell epitope; T cell receptor;
 KM down regulation; immune response; allergen; immunoglobulin E;
 KM sensitivity; cat protein allergen; human.
 XX
 OS Homo sapiens.
 XX
 PN US6019972-A.
 PN
 XX 01-FEB-2000.
 PD
 XX 02-SEP-1994; 94US-0300928.
 PF

```

XX 03-NOV-1989; 89US-0431565.
PR 28-FEB-1991; 91US-0662276.
PR 13-DEC-1991; 91US-0807529.
PR 25-MAR-1992; 92US-0857311.
PR 15-MAY-1992; 92US-0884718.
PR 15-JAN-1993; 93US-0006116.
XX
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
XX
PI Garman RD, Greenstein JL, Kuo M, Briner TJ, Morville M, Gelfer ML;
XX WPI: 2000-146862/13.
DR N-PSDB; AA288645.
XX
XX Peptides of human T cell reactive feline protein for treating
PT sensitivity to cat protein allergens comprise at least one T cell
PT epitope recognized by a T cell receptor specific for the human T cell
PT reactive feline protein.
XX
PS Example 25; Column 111-112; 105pp; English.
XX
CC This invention describes a novel peptide (I) of human T cell reactive
CC feline protein (TRFP) having at least one T cell epitope recognized
CC by a T cell receptor specific for the human T cell reactive feline
CC protein, the peptide consisting of at least 7-30 amino acids, and having
CC an amino acid sequence derived from an amino acid sequence comprising
CC 94, 96, 97, 109, or 111 residues, given in the specification. The
CC peptides down regulate the immune response to the allergen. The peptides
CC have reduced immunoglobulin E binding and reduce T cell responsiveness.
CC The peptide (I) is useful in compositions for treating sensitivity to a
CC cat protein allergen in a subject. AA51504-Y51516 represent peptides
CC which are used in the design of the oligonucleotides represented in AA288639-288650
CC human TRFP.
XX
SQ Sequence 4 AA;
XX
Query Match 78.9%; Score 15; DB 21; Length 4;
Best Local Similarity 50.0%; Pred. No. 7.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVPI 4
Db 1 ALPV 4
XX
RESUME 15
AA097418
ID AA097418 standard; Peptide; 4 AA.
XX
AC AA097418;
XX
XX 30-JUL-2002 (first entry)
XX
DE Synthetic tetrapeptide #2.
XX
XX Synthetic tetrapeptide; mimetic; inhibitor of apoptosis; IAP;
KW baculoviral IAP repeat; BIR; cell proliferative disease; cancer;
KW Alzheimer's disease; stroke; arthritis.
XX
XX Synthetic.
XX
XX OS
XX PN WO200226775-A2.
XX
XX PD 04-APR-2002.
XX
XX PF 28-SEP-2001; 2001WO-US30567.
XX
XX PR 29-SEP-2000; 2000US-236574P.
XX
XX PR 20-DEC-2000; 2000US-256830P.
XX
XX (UYPR-) UNIV PRINCETON.

```

```

XX Sh1 Y;
PI WPI: 2002-416473/44.
XX
XX Novel synthetic tetrapeptide, its mimetic or a compound useful for
PT rational drug design and in stimulating apoptosis in cell, binds an
PT inhibitor of apoptosis protein and relieves inhibition of caspase
XX
PS Claim 6; Page 43; 62pp; English.
XX
XX The present invention relates to a new synthetic tetrapeptide, its
XX non-peptide or partial peptide mimetic or a compound that binds an
XX inhibitor of apoptosis protein (IAP) and relieves IAP-mediated inhibition
XX of caspase activity, where the tetrapeptide binds a surface groove within
XX a baculoviral IAP repeat (BIR) domain of the IAP. The invention can be
XX used for stimulating apoptosis in a cultured cell and for screening for a
XX compound that binds an IAP at a surface groove within a BIR domain.
XX The synthetic tetrapeptide is a partial peptide or non-peptide mimetic
XX and is useful in making a drug suitable for treating cell proliferative
XX disease especially cancer in a mammal by promoting apoptosis in
XX proliferatively diseased cells. Agonists identified by the invention are
XX useful in treating cell proliferative diseases like cancer and
XX antagonists identified by the invention are useful in treating
XX Alzheimer's disease, stroke and arthritis. The present amino acid
XX sequence represents synthetic tetrapeptide #2 of the invention.
XX
SQ Sequence 4 AA;
XX
Query Match 78.9%; Score 15; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVP 3
Db 1 AVP 3

```

Search completed: February 16, 2003, 09:19:52
 Job time : 35 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:19:08 ; Search time 16 Seconds
(without alignments)
7.356 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 9158

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCtUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 17 | 89.5 | 4 | 6 | Patent No. 5496927-1 |
| 2 | 15 | 78.9 | 4 | 1 | US-07-807-529A-25 |
| 3 | 15 | 78.9 | 4 | 1 | US-08-002-202-7 |
| 4 | 15 | 78.9 | 4 | 1 | US-08-544-143A-20 |
| 5 | 15 | 78.9 | 4 | 2 | US-08-685-589A-33 |
| 6 | 15 | 78.9 | 4 | 3 | US-08-300-928C-74 |
| 7 | 15 | 78.9 | 4 | 3 | US-08-481-534-7 |
| 8 | 15 | 78.9 | 4 | 3 | US-08-430-944D-74 |
| 9 | 15 | 78.9 | 4 | 3 | US-08-430-014-74 |
| 10 | 15 | 78.9 | 4 | 3 | US-09-039-308A-16 |
| 11 | 15 | 78.9 | 4 | 3 | US-08-431-184-74 |
| 12 | 15 | 78.9 | 4 | 4 | US-09-578-303-23 |
| 13 | 15 | 78.9 | 4 | 1 | US-07-943-848-1 |
| 14 | 15 | 78.9 | 4 | 1 | US-08-155-331-18 |
| 15 | 15 | 78.9 | 4 | 1 | US-08-323-418-2 |
| 16 | 15 | 78.9 | 4 | 1 | US-08-323-418-3 |
| 17 | 15 | 78.9 | 4 | 1 | US-08-462-456-3 |
| 18 | 15 | 78.9 | 4 | 1 | US-08-002-202-15 |
| 19 | 15 | 78.9 | 4 | 1 | US-08-276-936A-1 |
| 20 | 15 | 78.9 | 4 | 1 | US-08-345-820B-4 |
| 21 | 15 | 78.9 | 4 | 1 | US-08-544-143A-16 |
| 22 | 15 | 78.9 | 4 | 1 | US-08-424-022-18 |
| 23 | 15 | 78.9 | 4 | 1 | US-08-224-868-12 |
| 24 | 15 | 78.9 | 4 | 1 | US-08-483-801-2 |
| 25 | 15 | 78.9 | 4 | 1 | US-08-483-801-3 |
| 26 | 15 | 78.9 | 4 | 1 | US-08-439-534-9 |
| 27 | 15 | 78.9 | 4 | 1 | US-08-777-208-4 |

| | | | | | | |
|----|----|------|---|---|--------------------|--------------------|
| 28 | 14 | 73.7 | 4 | 1 | US-08-777-208-5 | Sequence 5, Appl1 |
| 29 | 14 | 73.7 | 4 | 2 | US-08-722-268-1 | Sequence 1, Appl1 |
| 30 | 14 | 73.7 | 4 | 2 | US-08-394-600B-28 | Sequence 28, Appl1 |
| 31 | 14 | 73.7 | 4 | 2 | US-08-439-005-9 | Sequence 9, Appl1 |
| 32 | 14 | 73.7 | 4 | 2 | US-08-698-575E-4 | Sequence 4, Appl1 |
| 33 | 14 | 73.7 | 4 | 2 | US-08-463-076E-334 | Sequence 15, App |
| 34 | 14 | 73.7 | 4 | 2 | US-08-685-589A-15 | Sequence 28, Appl |
| 35 | 14 | 73.7 | 4 | 2 | US-08-685-589A-28 | Sequence 78, Appl |
| 36 | 14 | 73.7 | 4 | 2 | US-08-685-589A-78 | Sequence 18, Appl |
| 37 | 14 | 73.7 | 4 | 2 | US-08-424-017B-18 | Sequence 1, Appl1 |
| 38 | 14 | 73.7 | 4 | 2 | US-08-825-193-1 | Sequence 2, Appl1 |
| 39 | 14 | 73.7 | 4 | 2 | US-08-907-840A-2 | Sequence 1, Appl1 |
| 40 | 14 | 73.7 | 4 | 2 | US-08-985-056-1 | Sequence 1, Appl1 |
| 41 | 14 | 73.7 | 4 | 3 | US-08-984-884-1 | Sequence 1, Appl1 |
| 42 | 14 | 73.7 | 4 | 3 | US-09-090-046-1 | Sequence 1, Appl1 |
| 43 | 14 | 73.7 | 4 | 3 | US-09-090-274-1 | Sequence 1, Appl1 |
| 44 | 14 | 73.7 | 4 | 3 | US-08-984-881-1 | Sequence 1, Appl1 |
| 45 | 14 | 73.7 | 4 | 3 | US-09-069-823-4 | Sequence 4, Appl1 |

ALIGNMENTS

RESULT 1
5496927-1
Patent No. 5496927
; APPLICANT: KOEB, H.MICHAEL, BURKHART, JOSEPH P.; JUNG, MICHEL J.;
; GERRARD, FRITZ E.; GIKOUX, EUGENE L.; NEISS, BERNHARD; SCHIRKLIN,
; DANIEL G.
; TITLE OF INVENTION: PEPTIDASE INHIBITORS
; NUMBER OF SEQUENCES: 26
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,847
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 102,522
; FILING DATE: 04-AUG-1993
; APPLICATION NUMBER: 980,141
; FILING DATE: 23-NOV-1992
; APPLICATION NUMBER: 540,033
; FILING DATE: 19-JUN-1990
; APPLICATION NUMBER: 372,162
; FILING DATE: 27-JUN-1989
; APPLICATION NUMBER: 267,758
; FILING DATE: 01-NOV-1988
; APPLICATION NUMBER: 874,721
; FILING DATE: 15-JUN-1986
; APPLICATION NUMBER: 697,987
; FILING DATE: 04-FEB-1985
; SEQ ID NO:1:
; LENGTH: 4
5496927-1
Query Match 89.5%; Score 17; DB 6; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
DB 1 AVPI 4
OY 1 AVPI 4
; APPLICANT: Rogers, Bruce L.
; APPLICANT: Morgenshtern, Jay
; APPLICANT: Bond, Julian F.
; APPLICANT: Garman, Richard D.
; APPLICANT: Greenstein, Julia L.
; APPLICANT: Kuo, Mei-Chang
; RESULT 2
US-07-807-529A-25
; Sequence 25, Application US/07807529A
; Patent No. 5547669
; GENERAL INFORMATION:
; APPLICANT: Rogers, Bruce L.
; APPLICANT: Morgenshtern, Jay
; APPLICANT: Bond, Julian F.
; APPLICANT: Garman, Richard D.
; APPLICANT: Greenstein, Julia L.
; APPLICANT: Kuo, Mei-Chang

APPLICANT: Morville, Malcolm
TITLE OF INVENTION: RECOMBINANT PEPTIDES
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMLOGIC PHARMACEUTICAL CORPORATION
STREET: One Kendall Square, Building 600
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/807,529A
FILING DATE: 19911213
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,276
FILING DATE: 28-FEB-1991
APPLICATION NUMBER: US 07/431,565
FILING DATE: 03-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Channing, Stacey L.
REGISTRATION NUMBER: 31,095
REFERENCE/DOCKET NUMBER: IPC-027/4m1-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 494-0060
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: Peptide
FRAGMENT TYPE: Internal
US-07-807-529A-25

Query Match 78.9%; Score 15; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
1:1
Db 1 ALPV 4

RESULT 3
US-08-002-202-7
Sequence 7, Application US/08002202
Patent No. 5604201
GENERAL INFORMATION:
APPLICANT: Thomas, Garry
APPLICANT: Anderson, Eric D
APPLICANT: Thomas, Laurel
APPLICANT: Haylick, Joel S
TITLE OF INVENTION: Methods and Reagents for Inhibiting
TITLE OF INVENTION: Furin Endoprotease
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti and Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/002,202
FILING DATE: 08-JAN-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: No. 5604201nan, Kevin E
REGISTRATION NUMBER: 35,30003
REFERENCE/DOCKET NUMBER: 92,448
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-002-202-7

Query Match 78.9%; Score 15; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
1:1
Db 1 ALPM 4

RESULT 4
US-08-544-143A-20
Sequence 20, Application US/08544143A
Patent No. 5646028
GENERAL INFORMATION:
APPLICANT: Leigh, Scott D.
TITLE OF INVENTION: NOVEL ALKALINE PROTEASE AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Majestic, Parsons, Slebert & Hsue
STREET: Four Embarcadero Center, Suite 1100
CITY: San Francisco
STATE: California
COUNTRY: U.S.A.
ZIP: 94111-4121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,143A
FILING DATE: 17-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: de Runtz, K. Alison
REGISTRATION NUMBER: 37,119
REFERENCE/DOCKET NUMBER: 0409,054US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-362-5556
TELEFAX: 415-362-5418
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-544-143A-20

Query Match 78.9%; Score 15; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVPI 4
| | |
Db 1 AAPI 4

```
RESULT 5
US-08-685-589A-33
; Sequence 33, Application US/08685589A
; Patent No. 5916872
; GENERAL INFORMATION:
; APPLICANT: Chang, Conway
; APPLICANT: Gu, Leo
; APPLICANT: Chen, Jie
; TITLE OF INVENTION: CYCLIC PEPTIDES HAVING BROAD
; TITLE OF INVENTION: SPECTRUM ANTIMICROBIAL ACTIVITY
; NUMBER OF SEQUENCES: 222
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,589A
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 8067-026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: NO. 5916872 Relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..4
; OTHER INFORMATION: /product= "Beta-turn"
US-08-685-589A-33

Query Match 78.9%; Score 15; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02145
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/300,928C
; FILING DATE: September 2, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/807,529
; FILING DATE: December 13, 1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: AMY E. MANDRAGOURAS
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: 002.60US(IMI-044)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-300-928C-74
```

Query Match 78.9%; Score 15; DB 3; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
| | |
Db 1 ALPV 4

```
RESULT 7
US-08-481-534-7
; Sequence 7, Application US/08481534
; Patent No. 6022855
; GENERAL INFORMATION:
; APPLICANT: Thomas, Gary
; APPLICANT: Anderson, Eric D
; APPLICANT: Thomas, Laurel
; APPLICANT: Hayflick, Joel S
; APPLICANT: Nelson, Jay
; APPLICANT: Stenglen, Stephan G
; TITLE OF INVENTION: Methods and Reagents for Inhibiting Furin
; TITLE OF INVENTION: Endoprotease
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/481,534
FILING DATE: 14-SEP-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6022855nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 92,448-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-481-534-7

Query Match 78.9%; Score 15; DB 3; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Indels 0; Gaps 0;

OY 1 AVPI 4
1:1:
Db 1 ALPM 4

RESULT 8
US-08-430-944D-74
Sequence 74, Application US/08430944D
Patent No. 6025162
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,944D
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-430-944D-74

Query Match 78.9%; Score 15; DB 3; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Indels 0; Gaps 0;

OY 1 AVPI 4
1:1:
Db 1 ALPV 4

RESULT 9
US-08-430-014-74
Sequence 74, Application US/08430014
Patent No. 6048962
GENERAL INFORMATION:
APPLICANT: GEETER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: MALDEN
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,014
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/300,928
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.6US(IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-430-014-74

Query Match 78.9%; Score 15; DB 3; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
1:1:
Db 1 ALPV 4

RESULT 10
US-09-039-308A-16
Sequence 16, Application US/09039308A
Patent No. 6069129
GENERAL INFORMATION:
APPLICANT: Sandberg, Lawrence; Roos, Phillip;
TITLE OF INVENTION: ELASTIN DERIVED COMPOSITION
AND METHOD OF

TITLE OF INVENTION: USING SAME
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED SMITH SHAW & MCCRAY, LLP
CITY: PO Box 488
STATE: Pittsburgh
COUNTRY: Pennsylvania
ZIP: 15230
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB
COMPUTER: Compaq
OPERATING SYSTEM: Microsoft Windows 95
SOFTWARE: Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,308A
FILING DATE: March 13, 1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Raymond A.
REGISTRATION NUMBER: 42,891
REFERENCE/DOCKET NUMBER: 97-489
TELECOMMUNICATION INFORMATION:
TELEPHONE: (412) 288-4192
TELEFAX: (412) 288-3300
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-308A-16

Query Match 78.9%; Score 15; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVP 3
DB 1 AVP 3

RESULT 11
US-08-431-184-74
Sequence 74, Application US/08431184
Patent No. 6120769
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,184
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994

ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-431-184-74

Query Match 78.9%; Score 15; DB 3; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 1 ALPV 4

RESULT 12
US-09-578-303-23
Sequence 23, Application US/09578303
Patent No. 6399759
GENERAL INFORMATION:
APPLICANT: Travis, James
APPLICANT: Whitworth, S. Troy
APPLICANT: Blum, Murray S.
TITLE OF INVENTION: Ant Proteases and Methods of Inhibition
FILE REFERENCE: 235,00150101
CURRENT APPLICATION NUMBER: US/09/578,303
CURRENT FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/136,331
PRIOR FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 23
LENGTH: 4
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: acetyl end cap
NAME/KEY: SITE
LOCATION: (4)
OTHER INFORMATION: chloromethyl-ketone end cap
US-09-578-303-23

Query Match 78.9%; Score 15; DB 4; Length 4;
Best Local Similarity 75.0%; Pred. No. 2e+05; 1; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 1; Indels 0;

QY 1 AVPI 4
DB 1 AVPI 4

RESULT 13
US-07-943-848-1
Sequence 1, Application US/07943848
Patent No. 5322839
GENERAL INFORMATION:
APPLICANT: Voegel, Rainer
APPLICANT: Stocker, Kurt
APPLICANT: Mueller, Christian

TITLE OF INVENTION: Protein Fraction For Cosmetic and
TITLE OF INVENTION: Dermatologic Care of the Skin
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/943,848
FILING DATE: 19920911
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Charles E.
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 3006-024
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= No. 5322839e
OTHER INFORMATION: /note= "Residue 1 has an amide terminal methoxy
OTHER INFORMATION: succinyl group (MeoSuc)"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /label= No. 5322839e
OTHER INFORMATION: /note= "Residue 4 has a carboxy terminal
OTHER INFORMATION: para-nitroaniline group (pNA)"
US-07-943-848-1
Query Match 73.7%; Score 14; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AVPI 4
1 1:
Db 1 AAPV 4
RESULT 14
US-08-155-331-18
Sequence 18, Application US/08155331
Patent No. 5441931
GENERAL INFORMATION:
APPLICANT: Foster, Donald C
APPLICANT: Sprecher, Cindy
TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA

COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/155,331
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,692
FILING DATE: 02-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-684
REFERENCE/DOCKET NUMBER: 92-21C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-547-8080 ext 322
TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1..4
OTHER INFORMATION: /label= Ala-1
OTHER INFORMATION: /note= "Amino-terminal alanine residue is capped
OTHER INFORMATION: with a methoxysuccinyl gr..."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1..4
OTHER INFORMATION: /label= Val-4
OTHER INFORMATION: /note= "Carboxyl-terminal valine residue is capped
OTHER INFORMATION: with p-nitroanilide"
US-08-155-331-18
Query Match 73.7%; Score 14; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AVPI 4
1 1:
Db 1 AAPV 4
RESULT 15
US-08-323-418-2
Sequence 2, Application US/08323418
Patent No. 5478811
GENERAL INFORMATION:
APPLICANT: Peet, No. 5478811ton P
APPLICANT: Angelastro, Michael R
TITLE OF INVENTION: NO. 5478811e1 Orally-Active Elastase Inhibitors
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marion Merrell Dow Inc.
STREET: 2110 East Galbraith Rd.
CITY: Cincinnati P. O. Box 156300
STATE: Ohio
COUNTRY: USA
ZIP: 45215-6300
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/323,418
/ FILING DATE:
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/127,966
/ FILING DATE:
/ APPLICATION NUMBER: US/07/918,561
/ FILING DATE:
/ APPLICATION NUMBER: US 07/748,607
/ FILING DATE: 22-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Nesbitt, Stephen L.
/ REGISTRATION NUMBER: 28,981
/ REFERENCE/DOCKET NUMBER: M01627A US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (513) 948-7965
/ TELEFAX: (513) 948-7961
/ TELEX: 214320
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 1
/ OTHER INFORMATION: /note="morpholino carbonyl
/ OTHER INFORMATION: protected"
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 4
/ OTHER INFORMATION: /note="terminal OH is replaced by
/ OTHER INFORMATION: a perfluoroethyl group"
/ US-08-323-418-2

Query Match 73.7%; Score 14; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVPI 4
/ 1:
DB 1 AAPV 4
```

Search completed: February 16, 2003, 09:21:59
Job time : 16 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:21:03 ; Search time 10 seconds
(without alignments)
10.220 Million cell updates/sec

Title: US-09-939-293-13
Perfect score: 19
Sequence: 1 AVPI 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 1770

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCr_NEM_PUB pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCrUS_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEM_PUB pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 19 | 100.0 | 4 | 9 | US-10-068-569-21 |
| 2 | 19 | 100.0 | 4 | 9 | US-09-965-967-1 |
| 3 | 19 | 100.0 | 4 | 10 | US-09-939-293-13 |
| 4 | 15 | 78.9 | 4 | 9 | US-10-068-569-15 |
| 5 | 15 | 78.9 | 4 | 9 | US-09-965-967-4 |
| 6 | 15 | 78.9 | 4 | 9 | US-09-965-967-6 |
| 7 | 15 | 78.9 | 4 | 9 | US-09-965-967-24 |
| 8 | 15 | 78.9 | 4 | 10 | US-09-851-271A-11 |
| 9 | 15 | 78.9 | 4 | 10 | US-09-939-293-16 |
| 10 | 15 | 78.9 | 4 | 10 | US-09-939-293-17 |
| 11 | 14 | 73.7 | 4 | 9 | US-10-041-006A-11 |
| 12 | 14 | 73.7 | 4 | 9 | US-09-965-967-7 |
| 13 | 14 | 73.7 | 4 | 10 | US-09-842-543-1 |
| 14 | 14 | 73.7 | 4 | 10 | US-09-757-908A-14 |
| 15 | 14 | 73.7 | 4 | 10 | US-09-803-126-15 |
| 16 | 14 | 73.7 | 4 | 10 | US-09-117-380B-1 |
| 17 | 14 | 73.7 | 4 | 10 | US-09-939-293-18 |
| 18 | 14 | 73.7 | 4 | 12 | US-10-040-655-11 |
| 19 | 14 | 73.7 | 4 | 12 | US-10-154-507-6 |

| | | | | | | |
|----|----|------|---|----|--------------------|-------------------|
| 20 | 14 | 73.7 | 4 | 12 | US-10-154-507-13 | Sequence 13, Appl |
| 21 | 14 | 73.7 | 4 | 12 | US-10-154-507-14 | Sequence 14, Appl |
| 22 | 14 | 73.7 | 4 | 12 | US-10-154-507-15 | Sequence 15, Appl |
| 23 | 14 | 73.7 | 4 | 12 | US-10-154-507-16 | Sequence 16, Appl |
| 24 | 13 | 68.4 | 4 | 10 | US-09-935-744-2 | Sequence 2, Appl1 |
| 25 | 13 | 68.4 | 4 | 12 | US-10-033-526-3 | Sequence 3, Appl1 |
| 26 | 12 | 63.2 | 4 | 10 | US-10-033-526-2 | Sequence 2, Appl1 |
| 27 | 11 | 57.9 | 3 | 10 | US-09-853-918-2 | Sequence 2, Appl1 |
| 28 | 11 | 57.9 | 3 | 10 | US-09-853-918-5 | Sequence 5, Appl1 |
| 29 | 11 | 57.9 | 4 | 8 | US-08-484-409-6 | Sequence 6, Appl1 |
| 30 | 11 | 57.9 | 4 | 8 | US-08-424-530B-538 | Sequence 538, App |
| 31 | 11 | 57.9 | 4 | 9 | US-09-742-096-29 | Sequence 29, Appl |
| 32 | 11 | 57.9 | 4 | 9 | US-10-068-569-28 | Sequence 28, Appl |
| 33 | 11 | 57.9 | 4 | 9 | US-10-041-006A-10 | Sequence 10, Appl |
| 34 | 11 | 57.9 | 4 | 9 | US-09-972-035A-5 | Sequence 5, Appl1 |
| 35 | 11 | 57.9 | 4 | 9 | US-09-972-035A-6 | Sequence 6, Appl1 |
| 36 | 11 | 57.9 | 4 | 9 | US-09-965-967-5 | Sequence 3, Appl1 |
| 37 | 11 | 57.9 | 4 | 9 | US-10-087-905-3 | Sequence 3, Appl1 |
| 38 | 11 | 57.9 | 4 | 10 | US-09-802-077-19 | Sequence 19, Appl |
| 39 | 11 | 57.9 | 4 | 10 | US-09-802-077-56 | Sequence 56, Appl |
| 40 | 11 | 57.9 | 4 | 10 | US-09-802-096-19 | Sequence 19, Appl |
| 41 | 11 | 57.9 | 4 | 10 | US-09-802-096-36 | Sequence 36, Appl |
| 42 | 11 | 57.9 | 4 | 10 | US-09-837-969A-16 | Sequence 16, Appl |
| 43 | 11 | 57.9 | 4 | 10 | US-09-837-969A-41 | Sequence 41, Appl |
| 44 | 11 | 57.9 | 4 | 10 | US-09-757-908A-15 | Sequence 15, Appl |
| 45 | 11 | 57.9 | 4 | 10 | US-09-995-631-1 | Sequence 1, Appl1 |

ALIGNMENTS

RESULT 1
US-10-068-569-21
Sequence 21, Application US/10068569
; Patent No. US20020160975A1
; GENERAL INFORMATION:
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; TITLE OF INVENTION: CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS
; FILE REFERENCE: 480140.475
; CURRENT APPLICATION NUMBER: US/10/068,569
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-068-569-21

Query Match 100.0%; Score 19; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 1 AVPI 4

RESULT 2
US-09-965-967-1
; Sequence 1, Application US/09965967
; Patent No. US20020177557A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Yifeng
; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
; FILE REFERENCE: PU-0031 (01-1739-1)
; CURRENT APPLICATION NUMBER: US/09/965,967
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,574
; PRIOR FILING DATE: 2000-09-29

;; PRIOR APPLICATION NUMBER: 60/256,830
;; PRIOR FILING DATE: 2000-12-20
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 4
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-967-1

Query Match 100.0%; Score 19; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 3
US-09-939-293-13
;; Sequence 13, Application US/09939293
;; Patent No. US20020132786A1
;; GENERAL INFORMATION:
;; APPLICANT: Alnemri, Emad S.
;; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
;; TITLE OF INVENTION: AND METHODS OF USING THE SAME
;; FILE REFERENCE: 480140,465
;; CURRENT APPLICATION NUMBER: US/09/939,293
;; CURRENT FILING DATE: 2001-08-24
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13
;; LENGTH: 4
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-939-293-13

Query Match 100.0%; Score 19; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 4
US-10-068-569-15
;; Sequence 15, Application US/10068569
;; Patent No. US20020160975A1
;; GENERAL INFORMATION:
;; APPLICANT: Srinivasula, Srinivasa M.
;; APPLICANT: Fernandes-Alnemri, Teresa
;; APPLICANT: Alnemri, Emad S.
;; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
;; TITLE OF INVENTION: CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS
;; FILE REFERENCE: 480140,475
;; CURRENT APPLICATION NUMBER: US/10/068,569
;; CURRENT FILING DATE: 2002-02-06
;; NUMBER OF SEQ ID NOS: 28
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 15
;; LENGTH: 4
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-10-068-569-15

Query Match 78.9%; Score 15; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVP 3

Db 1 AVP 3
|||

RESULT 5
US-09-965-967-4
;; Sequence 4, Application US/09965967
;; Patent No. US20020177557A1
;; GENERAL INFORMATION:
;; APPLICANT: Shi, Yigong
;; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
;; FILE REFERENCE: PU-0031 (01-1739-1)
;; CURRENT APPLICATION NUMBER: US/09/965,967
;; CURRENT FILING DATE: 2001-09-28
;; PRIOR APPLICATION NUMBER: 60/236,574
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/256,830
;; PRIOR FILING DATE: 2000-12-20
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 4
;; LENGTH: 4
;; TYPE: PRT
;; ORGANISM: Drosophila melanogaster
US-09-965-967-4

Query Match 78.9%; Score 15; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVP 3
|||
Db 1 AVP 3

RESULT 6
US-09-965-967-6
;; Sequence 6, Application US/09965967
;; Patent No. US20020177557A1
;; GENERAL INFORMATION:
;; APPLICANT: Shi, Yigong
;; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
;; FILE REFERENCE: PU-0031 (01-1739-1)
;; CURRENT APPLICATION NUMBER: US/09/965,967
;; CURRENT FILING DATE: 2001-09-28
;; PRIOR APPLICATION NUMBER: 60/236,574
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/256,830
;; PRIOR FILING DATE: 2000-12-20
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 6
;; LENGTH: 4
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Synthetic Peptide
US-09-965-967-6

Query Match 78.9%; Score 15; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVP 3
|||
Db 1 AVP 3

RESULT 7
US-09-965-967-24
;; Sequence 24, Application US/09965967
;; Patent No. US20020177557A1
;; GENERAL INFORMATION:


```
; APPLICANT: Shi, Yigong
; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
; FILE REFERENCE: PU-0031 (01-1739-1)
; CURRENT APPLICATION NUMBER: US/09/965,967
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,574
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/256,830
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 24
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-965-967-24

Query Match
Best Local Similarity 100.0%; DB 9; Length 4;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPI 4
   |||
Db 2 VPI 4

RESULT 8
US-09-851-271A-11
; Sequence 11, Application US/09851271A
; Patent No. US20020064824A1
; GENERAL INFORMATION:
; APPLICANT: Gendag Limited
; TITLE OF INVENTION: Screening System
; FILE REFERENCE: 674538-2003
; CURRENT APPLICATION NUMBER: US/09/851,271A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: PCT/GB99/03730
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: GB9824544.2
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 11
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)-(4)
; OTHER INFORMATION: smallest unit of stalling polypeptide sequence
; US-09-851-271A-11

Query Match
Best Local Similarity 78.9%; DB 10; Length 4;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVP 3
   |||
Db 2 AVP 4

RESULT 9
US-09-939-293-16
; Sequence 16, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemrl, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 16
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Drosophila sp.
; US-09-939-293-16

Query Match
Best Local Similarity 78.9%; DB 10; Length 4;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVP 3
   |||
Db 1 AVP 3

RESULT 10
US-09-939-293-17
; Sequence 17, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemrl, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 17
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-939-293-17

Query Match
Best Local Similarity 78.9%; DB 10; Length 4;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVP 3
   |||
Db 1 AVP 3

RESULT 11
US-10-041-006A-11
; Sequence 11, Application US/10041006A
; Patent No. US20020168754A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/041,006A
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 11
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide substrate
; US-10-041-006A-11

Query Match
Best Local Similarity 73.7%; DB 9; Length 4;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVPI 4
   | | |
```

Db 1 AAPV 4

```
RESULT 12
US-09-965-967-7
; Sequence 7, Application US/09965967
; Patent No. US2002017757A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Yigong
; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
; FILE REFERENCE: PU-0031 (01-1739-1)
; CURRENT APPLICATION NUMBER: US/09/965,967
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,574
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/256,830
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.1
; SEQ ID NO 7
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-09-965-967-7
```

Query Match 73.7%; Score 14; DB 9; Length 4;
Best Local Similarity 50.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVPI 4
1 I:
Db 1 AAPV 4

```
RESULT 13
US-09-842-543-1
; Sequence 1, Application US/09842543
; Patent No. US20020010315A1
; GENERAL INFORMATION:
; APPLICANT: Gyorkos, et al.
; TITLE OF INVENTION: PEPTOID AND NONPEPTOID CONTAINING
; FILE REFERENCE: 361239-016A
; CURRENT APPLICATION NUMBER: US/09/842,543
; CURRENT FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 09/325,512
; PRIOR FILING DATE: 1999-06-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence is a known commercially available
; OTHER INFORMATION: substrate for elastases.
; NAME/KEY: BLOCKED
; LOCATION: (1)...(1)
; OTHER INFORMATION: N-terminal Methoxysuccinyl group.
; NAME/KEY: BLOCKED
; LOCATION: (4)...(4)
; OTHER INFORMATION: C-terminal p-nitroaniline
US-09-842-543-1
```

Query Match 73.7%; Score 14; DB 10; Length 4;
Best Local Similarity 50.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVPI 4
1 I:
Db 1 AAPV 4

```
RESULT 14
US-09-757-908A-14
; Sequence 14, Application US/09757908A
; Patent No. US20020052468A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; TITLE OF INVENTION: Disulfide Core Polypeptides
; FILE REFERENCE: 98-1301
; CURRENT APPLICATION NUMBER: US/09/757,908A
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: US 09/326,039
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: US 60/088,136
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-908A-14
```

Query Match 73.7%; Score 14; DB 10; Length 4;
Best Local Similarity 50.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVPI 4
1 I:
Db 1 AAPV 4

```
RESULT 15
US-09-803-126-15
; Sequence 15, Application US/09803126
; Patent No. US20020099190A1
; GENERAL INFORMATION:
; APPLICANT: Brooks, Alan R.
; APPLICANT: Deng, Gary G.
; APPLICANT: Rubany, Gabor M.
; TITLE OF INVENTION: Estrogen-Regulated Unconventional Myosin-Related
; FILE REFERENCE: 015303-000310US
; CURRENT APPLICATION NUMBER: US/09/803,126
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/188,488
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentlin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Consensus
US-09-803-126-15
```

Query Match 73.7%; Score 14; DB 10; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.2e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPI 4
1 I:
Db 1 VPV 3

Search completed: February 16, 2003, 09:25:24
Job time : 11 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:19:58 ; Search time 132 Seconds

(without alignments)
19.537 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 46895

Minimum DB seq length: 0

Maximum DB seq length: 4

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCrus_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 19 | 100.0 | 4 | 1 | PCT-US02-03553-21 |
| 2 | 19 | 100.0 | 4 | 1 | PCT-US02-17342-3 |
| 3 | 19 | 100.0 | 4 | 1 | PCT-US02-17342-34 |
| 4 | 19 | 100.0 | 4 | 1 | PCT-US02-17342-66 |
| 5 | 19 | 100.0 | 4 | 1 | PCT-US02-17342-68 |
| 6 | 19 | 100.0 | 4 | 20 | US-09-645-075-5 |

| | | | | | | |
|----|----|-------|---|----|---------------------|-------------------|
| 7 | 19 | 100.0 | 4 | 20 | US-09-687-549-14 | Sequence 14, Appl |
| 8 | 19 | 100.0 | 4 | 23 | US-09-939-293-13 | Sequence 13, Appl |
| 9 | 19 | 100.0 | 4 | 23 | US-09-965-967-1 | Sequence 1, Appl |
| 10 | 19 | 100.0 | 4 | 24 | US-10-068-569-21 | Sequence 21, Appl |
| 11 | 18 | 94.7 | 4 | 1 | PCT-US01-50075A-109 | Sequence 109, App |
| 12 | 18 | 94.7 | 4 | 1 | PCT-US02-17342-14 | Sequence 14, Appl |
| 13 | 18 | 94.7 | 4 | 1 | PCT-US02-17342-22 | Sequence 22, Appl |
| 14 | 17 | 89.5 | 4 | 1 | PCT-US02-17342-19 | Sequence 19, Appl |
| 15 | 16 | 84.2 | 4 | 1 | PCT-US02-17342-12 | Sequence 12, Appl |
| 16 | 16 | 84.2 | 4 | 1 | PCT-US02-17342-27 | Sequence 27, Appl |
| 17 | 16 | 84.2 | 4 | 1 | PCT-US02-17342-47 | Sequence 47, Appl |
| 18 | 16 | 84.2 | 4 | 1 | PCT-US02-17342-51 | Sequence 51, Appl |
| 19 | 15 | 78.9 | 4 | 1 | PCT-US02-03553-15 | Sequence 15, Appl |
| 20 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-1 | Sequence 1, Appl |
| 21 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-2 | Sequence 2, Appl |
| 22 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-4 | Sequence 4, Appl |
| 23 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-6 | Sequence 6, Appl |
| 24 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-9 | Sequence 9, Appl |
| 25 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-11 | Sequence 11, Appl |
| 26 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-13 | Sequence 13, Appl |
| 27 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-15 | Sequence 15, Appl |
| 28 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-20 | Sequence 20, Appl |
| 29 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-21 | Sequence 21, Appl |
| 30 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-23 | Sequence 23, Appl |
| 31 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-24 | Sequence 24, Appl |
| 32 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-25 | Sequence 25, Appl |
| 33 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-26 | Sequence 26, Appl |
| 34 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-28 | Sequence 28, Appl |
| 35 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-29 | Sequence 29, Appl |
| 36 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-30 | Sequence 30, Appl |
| 37 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-31 | Sequence 31, Appl |
| 38 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-33 | Sequence 33, Appl |
| 39 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-35 | Sequence 35, Appl |
| 40 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-55 | Sequence 55, Appl |
| 41 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-59 | Sequence 59, Appl |
| 42 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-63 | Sequence 63, Appl |
| 43 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-64 | Sequence 64, Appl |
| 44 | 15 | 78.9 | 4 | 1 | PCT-US02-17342-65 | Sequence 65, Appl |
| 45 | 15 | 78.9 | 4 | 1 | PCT-US02-22658-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1
PCT-US02-03553-21
; Sequence 21, Application PC/TUS0203553
; GENERAL INFORMATION:
; APPLICANT: Thomas Jefferson University
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; FILE REFERENCE: CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS
; CURRENT APPLICATION NUMBER: PCT/US02/03553
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-03553-21

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
1111
DB 1 AVPI 4

RESULT 2
PCT-US02-17342-3
; Sequence 3, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that B
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCT/US02/17342
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
PCT-US02-17342-3

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 3
PCT-US02-17342-34
; Sequence 34, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that B
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCT/US02/17342
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
; NAME/KEY: misc-feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: V is N-methylvaline
PCT-US02-17342-34

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 4
PCT-US02-17342-66
; Sequence 66, Application PC/TUS0217342
; GENERAL INFORMATION:

; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCT/US02/17342
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
; NAME/KEY: misc-feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: I is N-Methylisoleucine
PCT-US02-17342-66

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 5
PCT-US02-17342-68
; Sequence 68, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCT/US02/17342
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
; NAME/KEY: misc-feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: V is N-methylvaline
PCT-US02-17342-68

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 6
US-09-645-075-5
; Sequence 5, Application US/09645075
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong

```
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Apoptotic Compounds
; FILE REFERENCE: UTS0732
; CURRENT APPLICATION NUMBER: US/09/645,075
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-645-075-5

Query Match          100.0%; Score 19; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 7
US-09-687-549-14
; Sequence 14, Application US/09687549
; GENERAL INFORMATION:
; APPLICANT: Fesik, Stephen W.
; APPLICANT: Betz, Stephen F.
; APPLICANT: Meadows, Robert P.
; APPLICANT: Liu, Zhihong
; APPLICANT: Olejniczak, Edward T.
; APPLICANT: Sun, Chaohong
; TITLE OF INVENTION: PEPTIDES FROM SMAC (DIABLO) AND METHODS
; FILE REFERENCE: 6742.US.O1
; CURRENT APPLICATION NUMBER: US/09/687,549
; CURRENT FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide derived from smac
US-09-687-549-14

Query Match          100.0%; Score 19; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 8
US-09-939-293-13
; Sequence 13, Application US/09939293
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 4
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-13

Query Match          100.0%; Score 19; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 9
US-09-965-967-1
; Sequence 1, Application US/09965967
; GENERAL INFORMATION:
; APPLICANT: Shi, Yigong
; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
; FILE REFERENCE: PU-0031 (01-1739-1)
; CURRENT APPLICATION NUMBER: US/09/965,967
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,574
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/256,830
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-967-1

Query Match          100.0%; Score 19; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 10
US-10-068-569-21
; Sequence 21, Application US/10068569
; GENERAL INFORMATION:
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; FILE REFERENCE: 480140.475
; CURRENT APPLICATION NUMBER: US/10/068,569
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-068-569-21

Query Match          100.0%; Score 19; DB 24; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 11
```

```
PCN-US01-50075A-109
; Sequence 109, Application PC/TUS0150075A
; GENERAL INFORMATION:
; APPLICANT: Palatin Technologies, Inc.
; APPLICANT: Sharma, Shudh D.
; APPLICANT: Shi, Yi-Qun
; TITLE OF INVENTION: Identification of Target-Specific Folding Sites in Peptides and F
; FILE REFERENCE: 70025-PCN-14
; CURRENT APPLICATION NUMBER: PCN/TUS01/50075A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/256,842
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/304,835
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US 60/327,835
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 171
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 109
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C-terminus sequence for prion disease treatment metalloprotein
; OTHER INFORMATION: brary
PCN-US01-50075A-109
```

```
Query Match          94.7%; Score 18; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AVPI 4
       111:
Db       1 AVPV 4
```

```
RESULT 12
PCN-US02-17342-14
; Sequence 14, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that B
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCN/TUS02/17342
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
PCN-US02-17342-14
```

```
Query Match          94.7%; Score 18; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AVPI 4
       111:
Db       1 AVPI 4
```

```
RESULT 13
PCN-US02-17342-22
; Sequence 22, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
```

```
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCN/TUS02/17342
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
PCN-US02-17342-22
```

```
Query Match          94.7%; Score 18; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AVPI 4
       111:
Db       1 AVPV 4
```

```
RESULT 14
PCN-US02-17342-19
; Sequence 19, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCN/TUS02/17342
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
PCN-US02-17342-19
```

```
Query Match          89.5%; Score 17; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 AVPI 4
       111:
Db       1 AVPL 4
```

```
RESULT 15
PCN-US02-17342-12
; Sequence 12, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCN/TUS02/17342
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
```

SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 4
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic tetrapeptide sequence
PCT-US02-17342-12

Query Match 84.28; Score 16; DB 1; Length 4;
Best Local Similarity 75.08; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPI 4
|:|
Db 1 ALPI 4

Search completed: February 16, 2003, 09:24:24
Job time : 132 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:20:18 : Search time 23 Seconds

(without alignments)
14.693 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 429898 segs, 84487048 residues

Total number of hits satisfying chosen parameters: 1996

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/prodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 19 | 100.0 | 4 | 1 PCT-US02-37577-4 | Sequence 4, Appl |
| 2 | 19 | 100.0 | 4 | US-10-302-811-4 | Sequence 4, Appl |
| 3 | 15 | 78.9 | 3 | 1 PCT-US02-14358-21 | Sequence 21, Appl |
| 4 | 15 | 78.9 | 3 | US-10-141-531-21 | Sequence 21, Appl |
| 5 | 15 | 78.9 | 4 | US-09-580-893C-16 | Sequence 16, Appl |
| 6 | 15 | 78.9 | 4 | US-09-580-893D-16 | Sequence 16, Appl |
| 7 | 15 | 78.9 | 4 | US-09-580-110E-16 | Sequence 16, Appl |
| 8 | 15 | 78.9 | 4 | US-10-029-206A-13 | Sequence 13, Appl |
| 9 | 15 | 78.9 | 4 | US-10-028-075B-13 | Sequence 13, Appl |
| 10 | 15 | 73.7 | 4 | US-09-581-810A-1 | Sequence 1, Appl |
| 11 | 14 | 73.7 | 4 | US-09-580-893C-20 | Sequence 20, Appl |
| 12 | 14 | 73.7 | 4 | US-09-925-841-1 | Sequence 1, Appl |
| 13 | 14 | 73.7 | 4 | US-09-580-893D-20 | Sequence 20, Appl |
| 14 | 14 | 73.7 | 4 | US-09-780-503B-1 | Sequence 1, Appl |
| 15 | 14 | 73.7 | 4 | US-09-725-841-1 | Sequence 1, Appl |
| 16 | 14 | 73.7 | 4 | US-09-580-110E-20 | Sequence 20, Appl |
| 17 | 14 | 73.7 | 4 | US-10-259-609-29 | Sequence 29, Appl |
| 18 | 14 | 73.7 | 4 | US-10-190-082-224 | Sequence 224, Appl |
| 19 | 13 | 68.4 | 4 | US-09-402-923A-48 | Sequence 48, Appl |
| 20 | 12 | 63.2 | 4 | US-10-029-206A-76 | Sequence 76, Appl |
| 21 | 12 | 63.2 | 4 | US-10-028-075B-76 | Sequence 76, Appl |
| 22 | 11 | 57.9 | 4 | 1 PCT-US02-08146-5 | Sequence 5, Appl |
| 23 | 11 | 57.9 | 4 | 1 PCT-US02-08146-6 | Sequence 6, Appl |
| 24 | 11 | 57.9 | 4 | 1 PCT-US01-32150-9 | Sequence 9, Appl |
| 25 | 11 | 57.9 | 4 | 1 PCT-US01-32150-129 | Sequence 129, Appl |
| 26 | 11 | 57.9 | 4 | 4 US-08-761-902F-72 | Sequence 72, Appl |

ALIGNMENTS

| | | | | | |
|----|----|------|---|----------------------|--------------------|
| 27 | 11 | 57.9 | 4 | 5 US-09-517-466D-257 | Sequence 257, Appl |
| 28 | 11 | 57.9 | 4 | 5 US-09-580-893C-15 | Sequence 15, Appl |
| 29 | 11 | 57.9 | 4 | 5 US-09-580-893C-17 | Sequence 17, Appl |
| 30 | 11 | 57.9 | 4 | 5 US-09-580-893C-28 | Sequence 28, Appl |
| 31 | 11 | 57.9 | 4 | 5 US-09-580-893C-42 | Sequence 42, Appl |
| 32 | 11 | 57.9 | 4 | 5 US-09-580-893C-43 | Sequence 43, Appl |
| 33 | 11 | 57.9 | 4 | 5 US-09-580-893C-55 | Sequence 55, Appl |
| 34 | 11 | 57.9 | 4 | 5 US-09-580-893D-15 | Sequence 15, Appl |
| 35 | 11 | 57.9 | 4 | 5 US-09-580-893D-17 | Sequence 17, Appl |
| 36 | 11 | 57.9 | 4 | 5 US-09-580-893D-28 | Sequence 28, Appl |
| 37 | 11 | 57.9 | 4 | 5 US-09-580-893D-42 | Sequence 42, Appl |
| 38 | 11 | 57.9 | 4 | 5 US-09-580-893D-43 | Sequence 43, Appl |
| 39 | 11 | 57.9 | 4 | 5 US-09-580-893D-55 | Sequence 55, Appl |
| 40 | 11 | 57.9 | 4 | 5 US-09-721-507-1 | Sequence 1, Appl |
| 41 | 11 | 57.9 | 4 | 5 US-09-368-670C-7 | Sequence 7, Appl |
| 42 | 11 | 57.9 | 4 | 5 US-09-580-110E-15 | Sequence 15, Appl |
| 43 | 11 | 57.9 | 4 | 5 US-09-580-110E-17 | Sequence 17, Appl |
| 44 | 11 | 57.9 | 4 | 5 US-09-580-110E-28 | Sequence 28, Appl |
| 45 | 11 | 57.9 | 4 | 5 US-09-580-110E-43 | Sequence 43, Appl |

RESULT 1
PCT-US02-37577-4
Sequence 4, Application PC/TUS0237577
GENERAL INFORMATION:
APPLICANT: The Burnham Institute
APPLICANT: Reed, John C.
APPLICANT: Houghten, Richard A.
APPLICANT: Neftzi, Adel
APPLICANT: Ostresh, John M.
APPLICANT: Pinilla, Clemencia
APPLICANT: Pinilla, Kate
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Derepression of IAP-Inhibited Caspase
FILE REFERENCE: RP-LJ 5449
CURRENT APPLICATION NUMBER: PCT/US02/37577
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/331,957
PRIOR FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 4
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
PCT-US02-37577-4
Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVPI 4
Db 1 AVPI 4
RESULT 2
US-10-302-811-4
Sequence 4, Application US/10302811
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Houghten, Richard A.
APPLICANT: Neftzi, Adel
APPLICANT: Ostresh, John M.
APPLICANT: Pinilla, Clemencia
APPLICANT: Pinilla, Kate
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Derepression of IAP-Inhibited Caspase

```

; FILE REFERENCE: P-LT 5504
; CURRENT APPLICATION NUMBER: US/10/302,811
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/331,957
; PRIOR FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-302-811-4
```

```

Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AVPI 4
        |||
Db       1 AVPI 4
```

```

RESULT 3
PCT-US02-14358-21
; Sequence 21, Application PC/TUS0214358
; GENERAL INFORMATION:
; APPLICANT: XENCOR
; APPLICANT: STINGENTA PARTICIPATIONS AG
; APPLICANT: Briggs, Steven P.
; APPLICANT: Dalma, Bipin K.
; APPLICANT: del Val, Greg
; APPLICANT: Desjarlais, John R.
; APPLICANT: Heifetz, Peter
; APPLICANT: Lugibuhl, Peter
; APPLICANT: Muchhal, Umesh
; FILE REFERENCE: FP-71457-2-PC/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: PCT/US02/14358
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: US 60/370,609
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/289,029
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US02-14358-21
```

```

Query Match
Best Local Similarity 78.9%; Score 15; DB 1; Length 3;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 VPI 4
        |||
Db       1 VPI 3
```

```

RESULT 4
US-10-141-531-21
; Sequence 21, Application US/10141531
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Dalma, Bipin K.
; APPLICANT: del Val, Greg
; APPLICANT: Desjarlais, John R.
; APPLICANT: Heifetz, Peter
; APPLICANT: Lugibuhl, Peter
; APPLICANT: Muchhal, Umesh
```

```

; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
; FILE REFERENCE: A-71457-2/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/141,531
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US 60/370,609
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/289,029
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-141-531-21
```

```

Query Match
Best Local Similarity 78.9%; Score 15; DB 6; Length 3;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 VPI 4
        |||
Db       1 VPI 3
```

```

RESULT 5
US-09-580-893C-16
; Sequence 16, Application US/09580893C
; GENERAL INFORMATION:
; APPLICANT: SANDBERG, LAWRENCE B
; APPLICANT: MITTS, THOMAS F
; APPLICANT: JIMENEZ JR, FELIPE
; TITLE OF INVENTION: ASPARAGINE CONTAINING ELASTIN PEPTIDE ANALOGS
; FILE REFERENCE: 00-144-US
; CURRENT APPLICATION NUMBER: US/09/580,893C
; CURRENT FILING DATE: 2002-10-08
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 4
; TYPE: PRT
; ORGANISM: mammalian
US-09-580-893C-16
```

```

Query Match
Best Local Similarity 78.9%; Score 15; DB 5; Length 4;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AVP 3
        |||
Db       1 AVP 3
```

```

RESULT 6
US-09-580-893D-16
; Sequence 16, Application US/09580893D
; GENERAL INFORMATION:
; APPLICANT: SANDBERG, LAWRENCE B
; APPLICANT: MITTS, THOMAS F
; APPLICANT: JIMENEZ JR, FELIPE
; TITLE OF INVENTION: ASPARAGINE CONTAINING ELASTIN PEPTIDE ANALOGS
; FILE REFERENCE: 00-144-US
; CURRENT APPLICATION NUMBER: US/09/580,893D
; CURRENT FILING DATE: 2002-10-08
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 4
; TYPE: PRT
; ORGANISM: mammalian
US-09-580-893D-16
```

```

Query Match
Best Local Similarity 78.9%; Score 15; DB 5; Length 4;
```

Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 AVP 3
|||
DB 1 AVP 3

RESULT 7

US-09-580-110E-16
; Sequence 16, Application US/09580110E
; GENERAL INFORMATION:
; APPLICANT: Mits, Thomas F.
; APPLICANT: Sandberg, Lawrence B.
; TITLE OF INVENTION: ELASTIN PEPTIDE ANALOGS AND USES OF SAME IN COMBINATION WITH SKIN
; FILE REFERENCE: 25812-13
; CURRENT APPLICATION NUMBER: US/09/580,110E
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide
US-09-580-110E-16

Query Match 78.9%; Score 15; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 AVP 3
|||
DB 1 AVP 3

RESULT 8

US-10-029-206A-13
; Sequence 13, Application US/10029206A
; GENERAL INFORMATION:
; APPLICANT: Khan, Nisar A.
; APPLICANT: Benner, Robert
; TITLE OF INVENTION: Oligopeptide treatment of anthrax
; FILE REFERENCE: 2183-52220S
; CURRENT APPLICATION NUMBER: US/10/029,206A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/821,380
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 175
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: pdb/lftzv/lftzv-A
US-10-029-206A-13

Query Match 78.9%; Score 15; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 AVP 3
|||
DB 2 AVP 4

RESULT 9
US-10-028-075B-13
; Sequence 13, Application US/10028075B

; GENERAL INFORMATION:
; APPLICANT: Khan, Nisar A.
; APPLICANT: Benner, Robert
; TITLE OF INVENTION: Gene regulator
; FILE REFERENCE: 2183-52230S
; CURRENT APPLICATION NUMBER: US/10/028,075B
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EP 01203748.7
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 175
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: pdb/lftzv/lftzv-A
US-10-028-075B-13

Query Match 78.9%; Score 15; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 AVP 3
|||
DB 2 AVP 4

RESULT 10

US-09-581-810A-1
; Sequence 1, Application US/09581810A
; GENERAL INFORMATION:
; APPLICANT: Senol, Karim
; APPLICANT: Durans, Patrick
; APPLICANT: Colliet-Jouault, Sylvia
; APPLICANT: Singuin, Corinne
; APPLICANT: Jozefonvicz, Jacqueline
; APPLICANT: Letourneur, Didier
; APPLICANT: Blondin, Catherine
; APPLICANT: Gogly, Bruno
; APPLICANT: Pellat, Bernard
; TITLE OF INVENTION: USE OF PFCAN FOR REGULATING THE RECONSTRUCTION OF CONNECTIVE T
; FILE REFERENCE: 193108us-0-pct
; CURRENT APPLICATION NUMBER: US/09/581,810A
; CURRENT FILING DATE: 2000-11-06
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-Meo-Succ-
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (4)..(4)
; OTHER INFORMATION: -PA
US-09-581-810A-1

Query Match 73.7%; Score 14; DB 5; Length 4;
Best Local Similarity 50.0%; Pred. No. 3.5e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0;

QY 1 AVPI 4
|||
DB 1 AAPV 4

RESULT 11
US-09-580-893C-20
; Sequence 20, Application US/09580893C
; GENERAL INFORMATION:
; APPLICANT: SANDBERG, LAWRENCE B
; APPLICANT: MITTS, THOMAS F
; APPLICANT: JIMENEZ JR, FELIPE
; TITLE OF INVENTION: ASPARAGINE CONTAINING ELASTIN PEPTIDE ANALOGS
; FILE REFERENCE: 00-144-US
; CURRENT APPLICATION NUMBER: US/09/580,893C
; CURRENT FILING DATE: 2002-10-08
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 4
; TYPE: PRT
; ORGANISM: mammalian
US-09-580-893C-20

Query Match 73.7%; Score 14; DB 5; Length 4;
Best Local Similarity 66.7%; Pred. No. 3.5e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVP 3
1 1
Db 1 AIP 3

RESULT 12
US-09-925-841-1
; Sequence 1, Application US/09925841
; GENERAL INFORMATION:
; APPLICANT: Barriault, Dennis
; APPLICANT: Caruelle, Jean-Pierre
; APPLICANT: Amiri, Ahmed
; APPLICANT: Gautron, Jean
; TITLE OF INVENTION: Drug and Pharmaceutical Composition For the Treatment Of Lesions
; FILE REFERENCE: P606050U1
; CURRENT APPLICATION NUMBER: US/09/925,841
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 08/714,177
; PRIOR FILING DATE: 1996-12-20
; PRIOR APPLICATION NUMBER: PCT/FR95/00401
; PRIOR FILING DATE: 1995-03-29
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-925-841-1

Query Match 73.7%; Score 14; DB 5; Length 4;
Best Local Similarity 50.0%; Pred. No. 3.5e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AVPI 4
1 1
Db 1 AAPV 4

RESULT 13
US-09-580-893D-20
; Sequence 20, Application US/09580893D
; GENERAL INFORMATION:
; APPLICANT: SANDBERG, LAWRENCE B
; APPLICANT: MITTS, THOMAS F
; APPLICANT: JIMENEZ JR, FELIPE
; TITLE OF INVENTION: ASPARAGINE CONTAINING ELASTIN PEPTIDE ANALOGS
; FILE REFERENCE: 00-144-US

; CURRENT APPLICATION NUMBER: US/09/580,893D
; CURRENT FILING DATE: 2002-10-08
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 4
; TYPE: PRT
; ORGANISM: mammalian
US-09-580-893D-20

Query Match 73.7%; Score 14; DB 5; Length 4;
Best Local Similarity 66.7%; Pred. No. 3.5e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVP 3
1 1
Db 1 AIP 3

RESULT 14
US-09-780-503B-1
; Sequence 1, Application US/09780503B
; GENERAL INFORMATION:
; APPLICANT: Sethi, Sanjay
; TITLE OF INVENTION: A Method for Detecting Bacterial Exacerbations of Chronic Lung
; FILE REFERENCE: 11520.0228
; CURRENT APPLICATION NUMBER: US/09/780,503B
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,620
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: part of synthetic substrate for elastase
US-09-780-503B-1

Query Match 73.7%; Score 14; DB 5; Length 4;
Best Local Similarity 50.0%; Pred. No. 3.5e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AVPI 4
1 1
Db 1 AAPV 4

RESULT 15
US-09-725-841-1
; Sequence 1, Application US/09725841
; GENERAL INFORMATION:
; APPLICANT: BARRITOU, DENIS
; APPLICANT: CARUELLE, JEAN-PIERRE
; APPLICANT: AMIRI, AHMED
; APPLICANT: GAUTRON, JEAN
; TITLE OF INVENTION: DRUG AND PHARMACEUTICAL COMPOSITION FOR THE TREATMENT OF LESIO
; FILE REFERENCE: P606050U1
; CURRENT APPLICATION NUMBER: US/09/725,841
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 08/714,177
; PRIOR FILING DATE: 1996-12-20
; PRIOR APPLICATION NUMBER: PCT/FR95/00401
; PRIOR FILING DATE: 1995-03-29
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthesized

US-09-725-841-1

Query Match 73.78; Score 14; DB 5; Length 4;
Best Local Similarity 50.0%; Pred. No. 3.5e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 1 AAPV 4

Search completed: February 16, 2003, 09:25:00
Job time : 23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:18:28 : Search time 21 seconds
(Without alignments)
18.311 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 86

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 11 | 57.9 | 4 | 2 | 154357 |
| 2 | 9 | 47.4 | 4 | 2 | A32039 |
| 3 | 8 | 42.1 | 3 | 3 | I78890 |
| 4 | 8 | 42.1 | 4 | 2 | A27897 |
| 5 | 7 | 36.8 | 3 | 3 | RHTDRO |
| 6 | 7 | 36.8 | 3 | 3 | RHPCR |
| 7 | 7 | 36.8 | 3 | 3 | RHSHT |
| 8 | 7 | 36.8 | 3 | 3 | A33802 |
| 9 | 7 | 36.8 | 3 | 3 | A33391 |
| 10 | 7 | 36.8 | 4 | 2 | A02147 |
| 11 | 7 | 36.8 | 4 | 2 | PE0140 |
| 12 | 7 | 36.8 | 4 | 2 | S33508 |
| 13 | 7 | 36.8 | 4 | 2 | S17255 |
| 14 | 7 | 36.8 | 4 | 2 | A34626 |
| 15 | 7 | 36.8 | 4 | 2 | I51049 |
| 16 | 7 | 36.8 | 4 | 2 | PT0240 |
| 17 | 7 | 36.8 | 4 | 2 | PT0271 |
| 18 | 7 | 36.8 | 4 | 2 | PT0675 |
| 19 | 7 | 36.8 | 4 | 2 | PT0675 |
| 20 | 5 | 26.3 | 4 | 2 | S18401 |
| 21 | 5 | 26.3 | 4 | 2 | T30569 |
| 22 | 5 | 26.3 | 4 | 2 | PT0551 |
| 23 | 4 | 21.1 | 3 | 3 | PT0010 |
| 24 | 4 | 21.1 | 3 | 3 | PT0578 |
| 25 | 4 | 21.1 | 3 | 3 | PT0578 |
| 26 | 4 | 21.1 | 3 | 3 | PT0622 |
| 27 | 4 | 21.1 | 3 | 3 | S68328 |
| 28 | 4 | 21.1 | 3 | 3 | S68328 |
| 29 | 4 | 21.1 | 4 | 2 | PL0146 |

| | | | | | | |
|----|---|------|---|---|--------|--------------------|
| 30 | 4 | 21.1 | 4 | 2 | A37832 | phenol 2-monooxyge |
| 31 | 4 | 21.1 | 4 | 2 | A48360 | gamma subunit of p |
| 32 | 4 | 21.1 | 4 | 2 | A61300 | 22k superhelical D |
| 33 | 4 | 21.1 | 4 | 2 | S43014 | hypothetical prote |
| 34 | 4 | 21.1 | 4 | 2 | B43848 | cell surface adhes |
| 35 | 4 | 21.1 | 4 | 2 | I40804 | endoglucanase F - |
| 36 | 4 | 21.1 | 4 | 2 | T46627 | hypothetical prote |
| 37 | 4 | 21.1 | 4 | 2 | S09478 | globulin IV alpha |
| 38 | 4 | 21.1 | 4 | 2 | U01273 | neuropeptide Antho |
| 39 | 4 | 21.1 | 4 | 2 | A32480 | achatin-I - giant |
| 40 | 4 | 21.1 | 4 | 2 | S39390 | myosin-11ght-chain |
| 41 | 4 | 21.1 | 4 | 2 | I61883 | protamine PI - ora |
| 42 | 4 | 21.1 | 4 | 2 | S43959 | Ig mu chain V regi |
| 43 | 4 | 21.1 | 4 | 2 | E44823 | synaposomal-assoc |
| 44 | 4 | 21.1 | 4 | 2 | PT0696 | T-cell receptor be |
| 45 | 4 | 21.1 | 4 | 2 | PT0645 | T-cell receptor be |

ALIGNMENTS

RESULT 1

154357
schwannomin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C:Accession: 154357
R:Huyh, D.P.; Nechiporuk, T.; Pulst, S.
Hum. Mol. Genet. 3, 1075-1079, 1994
A:Title: Alternative transcripts in the mouse neurofibromatosis type 2 (NF2) gene ar
A:Reference number: 154357; MUID:95072570; PMID:7981675
A:Accession: 154357
A:Status: Preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-4 <RES>
A:Cross-references: GB:I28838; NID:9454836; PIDN:AA57150.1; PID:9601923
A:Gene: NF2

Query Match 57.9%; Score 11; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VP 3
DB 1 VP 2

RESULT 2

A32039
tyrosine-melanocyte-stimulating hormone release-inhibiting factor 1 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 18-Aug-2000
C:Accession: A32039
R:Horvath, A.; Kastin, A.J.
J. Biol. Chem. 264, 2173-2179, 1989
A:Title: Isolation of tyrosine-melanocyte-stimulating hormone release-inhibiting fac
A:Reference number: A32039; MUID:89123285; PMID:2563371
A:Accession: A32039
A:Molecule type: protein
A:Residues: 1-4 <HOR>
A:Experimental source: brain
C:Superfamily: unassigned animal peptides
C:Keywords: amidated carboxyl end
F:4/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 47.4%; Score 9; DB 2; Length 4;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PI 4
DB 2 PL 3

RESULT 3
178890
tyrosine protein kinase - mouse (fragment)

C:Species: Mus musculus (house mouse)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: 178890
R:Chow, L.M.; Davidson, D.; Fournel, M.; Gosselin, P.; Lemieux, S.; Lyu, M.S.; Kozak, C.
Oncogene 9, 3437-3448, 1994
A:Title: Two distinct protein isoforms are encoded by nt, a csk-related tyrosine protein
A:Reference number: 158407; MUID:95060800; PMID:7970703
A:Accession: 178890
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3 <RES>
A:Cross-references: GB:U33339; NID:g609536; PIDN:AAA64432.1; PID:g609538
A:Gene: p52ntk

Query Match 42.1%; Score 8; DB 3; Length 3;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VP 3
1 1
Db 1 WP 2

RESULT 4

A27897
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - Aspergillus phoenicis (fragment)
N:Alternate names: glucoamylase
C:Species: Aspergillus phoenicis
C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 06-Dec-1996
C:Accession: A27897
R:Inokuchi, N.; Takahashi, T.; Irie, M.
J. Biochem. 90, 1055-1067, 1981
A:Title: Purification and characterization of a minor glucoamylase from Aspergillus saitoi
A:Reference number: A27897; MUID:82075730; PMID:6796572
A:Note: Aspergillus saitoi
A:Accession: A27897
A:Molecule type: protein
A:Residues: 1-4 <IMO>
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 42.1%; Score 8; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AV 2
1 1
Db 1 AV 2

RESULT 5

RHPTDO
thyroliberin - Bombina orientalis
C:Species: Bombina orientalis
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A90919; A01415
R:Yasunara, T.; Nakajima, T.
Chem. Pharm. Bull. 23, 3301-3303, 1975
A:Title: Occurrence of Pyr-His-Pro-NH₂ in the frog skin.
A:Reference number: A90919; MUID:76138399; PMID:815011
A:Accession: A90919
A:Molecule type: protein
A:Residues: 1-3 <YAS>
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 36.8%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 P 3
1 1
Db 3 P 3

RESULT 6

RHPTCT
thyroliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A01415
R:Nair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.
Biochemistry 9, 1103-1106, 1970
A:Title: Structure of porcine thyrotropin releasing hormone.
A:Reference number: A90560; MUID:70136150; PMID:4984938
A:Accession: A01415
A:Molecule type: protein
A:Residues: 1-3 <NAI>
R:Boyer, J.; Enzmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.
Biochem. Biophys. Res. Commun. 37, 705-710, 1969
A:Title: The identity of chemical and hormonal properties of the thyrotropin releasin
A:Reference number: A90167; MUID:70039904; PMID:4982117
A:Contents: annotation
A:Note: biological activities and Rf values (in 17 chromatographic systems) of the sy
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 36.8%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 P 3
1 1
Db 3 P 3

RESULT 7

RHSHT
thyroliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A93750; A01415
R:Desiderio Jr., D.M.; Burgess, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.
Org. Mass Spectrom. 5, 221-228, 1971
A:Title: The elucidation of the primary structure of the hypothalamic thyroid stimula
A:Reference number: A93750
A:Accession: A93750
A:Molecule type: protein
A:Residues: 1-3 <DES>
R:Burgus, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.
Nature 226, 321-325, 1970
A:Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor
A:Reference number: A93161; MUID:70163386; PMID:485794
A:Contents: annotation
A:Note: physicochemical characteristics and biological activities of the natural and
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 36.8%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 P 3
1 1
Db 3 P 3

RESULT 8

A92971
Thyroliberin - eastern newt (tentative sequence)
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A92971; A01415
R:Grinn-Jorgensen, Y.; McKelvy, J.F.
J. Neurochem. 23, 471-478, 1974
A:Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) brain
A:Reference number: A92971; MUID:75035605; PMID:4214528
A:Accession: A92971
A:Molecule type: protein
A:Residues: 1-3 <GRI>
A:Note: a peptide with the chromatographic and electrophoretic characteristics of thyroliberin, or glutamic acid
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic F;3/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match

36.8%; Score 7; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 P 3

DB 3 P 3

RESULT 9

A33802
Thyrotropin-releasing hormone-like peptide - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A33802
R:Cockle, S.M.; Aitken, A.; Beg, F.; Smyth, D.G.
J. Biol. Chem. 264, 7788-7791, 1989
A:Title: A novel peptide, pyroglutamylglutamyproline amide, in the rabbit prostate comp

A:Reference number: A33802; MUID:89255196; PMID:2498305
A:Accession: A33802
A:Status: preliminary
A:Molecule type: protein

A:Residues: 1-3 <COC>
C:Superfamily: unassigned animal peptides

C:Keywords: amidated carboxyl end; pyroglutamic acid

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 36.8%; Score 7; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 P 3

DB 3 P 3

RESULT 10

A43391
TRH-like tripeptide - alfalfa
C:Species: Medicago sativa (alfalfa)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A43391
R:Lackey, D.B.
J. Biol. Chem. 267, 17508-17511, 1992
A:Title: Isolation and structural determination of a novel TRH-like tripeptide, pyroglu-

A:Reference number: A43391; MUID:92388092; PMID:1517203
A:Accession: A43391

A:Molecule type: protein

A:Residues: 1-3 <LAC>

C:Keywords: amidated carboxyl end; pyroglutamic acid

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match

36.8%; Score 7; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 P 3

DB 3 P 3

RESULT 11

A02147
phagocytosis-stimulating peptide (tuftsin) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 03-Feb-1994
C:Accession: A02147
R:Kishore, K.; Constantopoulos, A.; Satoh, P.S.; Najjar, V.A.
Biochem. Biophys. Res. Commun. 47, 172-179, 1972
A:Title: The characteristics, isolation and synthesis of the phagocytosis stimulating
A:Reference number: A02147; MUID:72187087; PMID:4112769
A:Accession: A02147
A:Molecule type: protein
A:Residues: 1-4 <NTS>
A:Note: a peptide having the same structure, physical properties, and biological acti
R:Fidalgo, B.V.; Najjar, V.A.
Biochemistry 6, 3386-3392, 1967
A:Reference number: A37502; MUID:68091045; PMID:4169272
A:Contents: annotation: immunoglobulin class
C:Comment: An IgG (called leucokinin) binds reversibly to the cell membrane of neutro
n is essential for maximum stimulation of the phagocytic activity of neutrophils.
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match

36.8%; Score 7; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 P 3

DB 3 P 3

RESULT 12

PL0140
carbon-monoxide dehydrogenase (EC 1.2.99.2) large chain - Pseudomonas carboxydohydrog
C:Species: Pseudomonas carboxydohydrog
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Apr-1993
C:Accession: PL0140
R:Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.
Arch. Microbiol. 152, 335-341, 1989
A:Title: Homology and distribution of CO dehydrogenase structural genes in carboxydot
A:Reference number: PL0138; MUID:90055678; PMID:2818128
A:Accession: PL0140

A:Molecule type: protein

A:Residues: 1-4 <KRA>

C:Comment: Carbon-monoxide dehydrogenase consists of three polypeptide chains: large,

C:Keywords: oxidoreductase

Query Match

36.8%; Score 7; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 P 3

DB 4 P 4

RESULT 13

S53508
starvation-induced ribonuclease - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999

Job time : 22 secs

C:Accession: S53508
 R:Koeck, M.; Loeffler, A.; Abel, S.; Glund, K.
 Plant Mol. Biol. 27, 477-485, 1995
 A:Title: cDNA structure and regulatory properties of a family of starvation-induced ribo
 A:Reference number: S53506; MUID:95201242; PMID:7894013
 A:Accession: S53508
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-4 <KOE>

Query Match 36.8%; Score 7; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 P 3
 |
 DB 2 P 2

RESULT 14

S17255
 ribosomal protein Yml1, mitochondrial, questionable - yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 A:Variety: strain 07173
 C:Date: 23-Apr-1993 #sequence_revision 14-Sep-1994 #text_change 09-May-1997
 C:Accession: S17255
 R:Grohmann, L.; Grack, H.R.; Kruff, V.; Choll, T.; Goldschmidt-Reisin, S.; Kitakawa, M.
 FEBS Lett. 284, 51-56, 1991
 A:Title: Extended N-terminal sequencing of proteins of the large ribosomal subunit from
 A:Reference number: S17255; MUID:91285106; PMID:2060626
 A:Accession: S17255
 A:Molecule type: protein
 A:Residues: 1-4 <GRO>
 A:Comment: A coding region for this protein could not be identified in the genome of Sac
 C:Genetics:
 A:Genome: nuclear
 C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 36.8%; Score 7; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 P 3
 |
 DB 4 P 4

RESULT 15

A34626
 RPCH-related neuropeptide - ferruginous spindle
 C:Species: Fusinus ferrugineus (ferruginous spindle)
 C:Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 31-Dec-1993
 C:Accession: A34626
 R:Kuroki, Y.; Kanda, T.; Kubota, I.; Fujisawa, Y.; Ikeda, T.; Miura, A.; Minamitake, Y.;
 Biochem. Biophys. Res. Commun. 167, 273-279, 1990
 A:Title: A molluscan neuropeptide related to the crustacean hormone, RPCH.
 A:Reference number: A34626; MUID:90179762; PMID:2310394
 A:Accession: A34626
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-4 <KUR>
 C:Keywords: neuropeptide

Query Match 36.8%; Score 7; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 P 3
 |
 DB 2 P 2

Search completed: February 16, 2003, 09:21:31

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:11:23 : Search time 10 Seconds

(without alignments)
16,591 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19
Sequence: 1 AVPI 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 18

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 7 | 36.8 | 3 | 1 | THYL_PIG |
| 2 | 7 | 36.8 | 4 | 1 | DCML_PSECH |
| 3 | 7 | 36.8 | 4 | 1 | RM01_YEAST |
| 4 | 7 | 36.8 | 4 | 1 | TYFT_HUMAN |
| 5 | 4 | 21.1 | 3 | 1 | LUXE_VIBRI |
| 6 | 4 | 21.1 | 4 | 1 | ACH1_ACHRU |
| 7 | 4 | 21.1 | 4 | 1 | DCMS_PSECH |
| 8 | 4 | 21.1 | 4 | 1 | EOS1_HUMAN |
| 9 | 4 | 21.1 | 4 | 1 | FFKA_AMEL |
| 10 | 4 | 21.1 | 4 | 1 | FYRI_AMEL |
| 11 | 2 | 10.5 | 4 | 1 | FAR3_HIRME |
| 12 | 2 | 10.5 | 4 | 1 | FLRF_HIRME |
| 13 | 2 | 10.5 | 4 | 1 | FLRN_AMEL |
| 14 | 1 | 5.3 | 4 | 1 | FARN_HIRME |
| 15 | 1 | 5.3 | 4 | 1 | FMRF_MACNI |
| 16 | 1 | 5.3 | 4 | 1 | OCP3_OCTMI |
| 17 | 0 | 0.0 | 3 | 1 | GRMA_HUMAN |
| 18 | 0 | 0.0 | 4 | 1 | OCPI_OCTMI |

ALIGNMENTS

RESULT 1
THYL_PIG
ID THYL_PIG STANDARD: PRT: 3 AA.
AC P01151;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
OS Sus scrofa (Pig),
OS Ovis aries (Sheep),

OS Bombina orientalis (Oriental fire-bellied toad), and
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823, 9940, 8346, 8316;
RN [1]
RP SEQUENCE.
RC SPECIES-Pig: TISSUE-Hypothalamus;
RX MEDLINE-70136150; PubMed-4984938;
RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
RT "Structure of porcine thyrotropin releasing hormone.";
RL Biochemistry 9:1103-1106(1970).
RN [2]
RP SYNTHESIS.
RC SPECIES-Pig;
RX MEDLINE-70039904; PubMed-4982117;
RA Boler J., Enzmann F., Folkers R., Bowers C.Y., Schally A.V.;
RT "The identity of chemical and hormonal properties of the thyrotropin
RT releasing hormone and pyroglutamyl-histidyl-proline amide.";
RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
RN [3]
RP SEQUENCE.
RC SPECIES-Sheep: TISSUE-Hypothalamus;
RA Desiderio D.M. Jr., Burgess R., Dunn T.F., Vale W., Guillemin R.,
RA Ward D.N.;
RT "The elucidation of the primary structure of the hypothalamic thyroid
RT stimulating hormone releasing factor of ovine origin by means of mass
RT spectrometry.";
RL Org. Mass Spectrom. 5:221-228(1971).
RN [4]
RP SYNTHESIS.
RC SPECIES-Sheep;
RX MEDLINE-70163386; PubMed-4985794;
RA Burgess R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
RA Guillemin R.;
RT "Characterization of ovine hypothalamic hypophysiotropic
RT TSH-releasing factor.";
RL Nature 226:321-325(1970).
RN [5]
RP SEQUENCE.
RC SPECIES-B.orientalis; TISSUE-Skin;
RX MEDLINE-76138399; PubMed-815011;
RA Yasuhara T., Nakajima T.;
RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin.";
RL Chem. Pharm. Bull. 23:3301-3303(1975).
RN [6]
RP SEQUENCE.
RC SPECIES-N.viridescens;
RX MEDLINE-75035605; PubMed-4214528;
RA Glimm-Joergensen Y., McKelvy J.F.;
RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
RT viridescens) brain in vitro. Isolation and characterization of
RT thyrotropin releasing factor.";
RL J. Neurochem. 23:471-478(1974).
CC -1- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH
CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/
CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.
DR PIR: A01415; RHPGT.
DR PIR: A93750; RSHSR.
DR PIR: A90919; RHTDRO.
DR PIR: A92971; A92971.
KW Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 3 3 AMIDATION.
SQ SEQUENCE 3 AA; 380 MW; 7761F6B000000000 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 3;
Best local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 P 3
Db 3 P 3

RESULT 2

DCML_PSECH STANDARD: PRT: 4 AA.
 AC P19916;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO dehydrogenase subunit L) (CO-PH L) (Fragment).
 GN Cntrl.
 OS Pseudomonas carboxydohydrogena.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Bradyrhizobium group.
 CC NCBI_TaxID=290;
 OX [1]
 RN SEQUENCE.
 RP MEDLINE=90055678; PubMed=2818128;
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
 RT "Homology and distribution of Co dehydrogenase structural genes in carboxydotrophic bacteria.";
 RL Arch. Microbiol. 152:335-341(1989).
 CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon dioxide.
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced acceptor.
 CC -1- COFACTOR: Molybdenum (molybdopterin).
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND SMALL.
 DR PIR: PLO140; PLO140.
 KW Oxidoreductase; Molybdenum.
 FT NON_TER
 SQ SEQUENCE 4 AA; 441 MW; 7761E876F0000000 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 P 3
 DB 4 P 4

RESULT 3

RM01_YEAST STANDARD: PRT: 4 AA.
 AC P36515;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Mitochondrial 60S ribosomal protein L1 (YmlL) (Fragment).
 GN MRPL.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=91285106; PubMed=2060626;
 RA Grohmann L., Graack H.-R., Kruff V., Choll T., Goldschmidt-Reisin S., Kitakawa M.;
 RT "Extended N-terminal sequencing of proteins of the large ribosomal subunit from yeast mitochondria.";
 RL FEBS Lett. 284:51-56(1991).
 DR PIR: S17255; S17255.
 DR SGD: I0002681; MRPL.
 KW Ribosomal protein; Mitochondrion.
 FT NON_TER
 SQ SEQUENCE 4 AA; 402 MW; 7771B2D5D0000000 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 P 3
 DB 4 P 4

RESULT 4

TUFT_HUMAN STANDARD: PRT: 4 AA.
 AC P01858;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phagocytosis-stimulating peptide (Tuftsin).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 OX [1]
 RN SEQUENCE.
 RP MEDLINE=72187087; PubMed=4112769;
 RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;
 RT "The physiological role of the lymphoid system. VI. The stimulatory effect of leucophilic gamma globulin (leucokinin) on the phagocytic activity of human polymorphonuclear leucocyte.";
 RL Biochemistry 6:3386-3392(1967).
 CC -1- MISCELLANEOUS: AN IGG (CALLED LECOKININ) BINDS REVERSIBLY TO THE CELL MEMBRANE OF NEUTROPHILS IN THE BLOOD. LECOKININASE ON THE MEMBRANE RELEASES THE ACTIVE PEPTIDE TUFTSIN FROM THE GAMMA CHAIN. TUFTSIN IS ESSENTIAL FOR MAXIMUM STIMULATION OF THE PHAGOCYTIC ACTIVITY OF NEUTROPHILS.
 DR PIR: A02147; A02147.
 DR MIM: 191150;
 SQ SEQUENCE 4 AA; 501 MW; 74176321C0000000 CRC64;

Query Match 36.8%; Score 7; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 P 3
 DB 3 P 3

RESULT 5

LUXE_VIBFI STANDARD: PRT: 3 AA.
 AC P24272;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Long-chain-fatty-acid-1-acyltransferin-component ligase (EC 6.2.1.19) (Acyl-LUXE).
 GN LuxE.
 OS Vibrio fischeri.
 CC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 CC NCBI_TaxID=668;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=91072226; PubMed=2254256;
 RA Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;
 RT "A new Vibrio fischeri lux gene precedes a bidirectional termination site for the lux operon.";
 RL J. Bacteriol. 172:6797-6802(1990).
 CC -1- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.

CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE
 CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS
 CC SUBSTRATE IN THE LIPIDASE-CATALYZED REACTION.
 CC
 CC -1- CATALYTIC ACTIVITY: ATP + an acid + protein -> AMP + diphosphate +
 CC an acyl-protein thioester.
 CC
 CC -1- PATHWAY: Biolumenecent fatty acid reduction system; second step.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: M62812; -; NOT_ANNOTATED_CDS.
 CC Luminescence; Ligase.
 FT NON_TER 1
 SQ SEQUENCE 3 AA: 374 MW: 6AA3303000000000 CRC64;
 Query March 21.1%; Score 4; DB 1; Length 3;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 1 4
 Db 1 1 1
 RESULT 6
 ACH1_ACHFU STANDARD; PRT; 4 AA.
 ID ACH1_ACHFU
 AC P35904;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Achatin-I.
 OS Achatina fulica (Giant African snail).
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 CC Achatinacea; Achatinidae; Achatina.
 OC NCB1_TaxID=6530;
 OX NCB1
 RN
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 RC STRAIN=Perussac; TISSUE=Ganglion;
 RX MEDLINE=89273551; PubMed=2597281;
 RA Kanatani Y., Minakata H., Kenmy P.T.M., Iwashita T., Watanabe K.,
 RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,
 RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;
 RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
 RT fulica Ferussac containing a D-amino acid residue";
 RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
 RN
 RN [2]
 RP CHARACTERIZATION.
 RC STRAIN=Perussac; TISSUE=Heart atrium;
 RX MEDLINE=91264856; PubMed=1675568;
 RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
 RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
 RT "Purification of achatin-I from the atria of the African giant snail,
 RT Achatina fulica, and its possible function.";
 RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
 RN
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE=93014529; PubMed=1399265;
 RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,
 RA Iwashita T., Nomoto K.;
 RT "Crystal structure and molecular conformation of achatin-I
 RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
 RT D-amino acid residue";
 RL Int. J. Pept. Protein Res. 39:258-264(1992).
 CC
 CC -1- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY
 CC AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY
 CC NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE
 CC HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.

DR PIR: A32480; A32480.
 CC Hormone; D-amino acid.
 FT MOD_RES 2
 SQ SEQUENCE 4 AA: 408 MW: 6AADD9C810000000 CRC64;
 Query Match 21.1%; Score 4; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 A 1
 Db 3 A 3
 RESULT 7
 DCMS_PSECH STANDARD; PRT; 4 AA.
 ID DCMS_PSECH
 AC P1918;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO
 DE dehydrogenase subunit S) (CO-DH S) (Fragment).
 DE COPS.
 OS Pseudomonas carboxydohydrogena.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Bradyrhizobium group.
 OC NCB1_TaxID=290;
 OX NCB1
 RN
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=90055678; PubMed=2818128;
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
 RT "Homology and distribution of CO dehydrogenase structural genes in
 RT carboxydophilic bacteria";
 RL Arch. Microbiol. 152:335-341(1989).
 CC
 CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
 CC dioxide.
 CC
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor -> CO(2) + reduced
 CC acceptor.
 CC
 CC -1- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
 CC
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 CC SMALL.
 DR PIR: P10146; P10146.
 KW Oxidoreductase; Iron-sulfur.
 FT NON_TER 4
 SQ SEQUENCE 4 AA: 420 MW: 6DD33DD6F0000000 CRC64;
 Query Match 21.1%; Score 4; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 A 1
 Db 2 A 2
 RESULT 8
 EOS1_HUMAN STANDARD; PRT; 4 AA.
 ID EOS1_HUMAN
 AC P02731;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Eosinophilolactac peptide.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCB1_TaxID=9606;
 RN
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=76078412; PubMed=1060093;
 RA Goetzl E.J., Austen K.F.;
 RT "Purification and synthesis of eosinophilolactac tetrapeptides of

```

RT human lung tissue: identification as eosinophil chemotactic factor of
RT anaphylaxis."
RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
CC -1- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
DR PIR; A03190; ETHOL.
FT VARIANT 1 1 V -> A (IN OTHER PEPTIDE).
FT SEQUENCE 4 AA: 390 MW: 6805B862A0000000 CRC64:
SQ /FTID=VAR.005201.

Query Match 21.1%; Score 4; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 V 2
Db 1 V 1

RESULT 9
FFKA_AMEL STANDARD; PRT; 4 AA.
AC P58705;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antho-Kamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;
OC Nynanthaeae; Actiniidae; Anthopleura.
OX NCBI_Taxid=6110;
RN [1]
RP SEQUENCE.
RX MEDLINE=9202852; Pubmed=1681803;
RA Notacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
RT "Isolation of L-3-phenylacetyl-Phe-Lys-Ala-NH2 (Antho-Kamide)," a
RT novel neuropeptide from sea anemones."
RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
RN [2]
RP FUNCTION.
RX MEDLINE=93391436; Pubmed=8397415;
RA McFarlane I.D., Hudman D., Notacker H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two
RT inhibitory neuropeptides, Antho-Kamide and Antho-Riamide."
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC -1- FUNCTION: Inhibits spontaneous contractions in several muscle
CC groups. May be involved in the expansion phase of feeding
CC behaviour in sea anemones.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Neuron-specific.
KM Neuropeptide; Amidation.
FT MOD_RES 1 1 L-3-PHENYLACTYL.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA: 512 MW: 6DD339C9A0000000 CRC64:

Query Match 21.1%; Score 4; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 A 1
Db 4 A 4

RESULT 10
FYRI_AMEL STANDARD; PRT; 4 AA.
AC P58706;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antho-Riamide I [Contains: Antho-Riamide II].
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;
OC Nynanthaeae; Actiniidae; Anthopleura.
OX NCBI_Taxid=6110;
RN [1]
RP SEQUENCE.
RX MEDLINE=92270459; Pubmed=1821096;
RA Notacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,
RA Grimmelikhuijzen C.J.P.;
RT "Isolation of two novel neuropeptides from sea anemones: the unusual,
RT biologically active L-3-phenylacetyl-Tyr-Arg-Ile-NH2 and its
RT des-phenylacetyl fragment Tyr-Arg-Ile-NH2."
RL Peptides 12:1165-1173(1991).
RN [2]
RP FUNCTION.
RX MEDLINE=93391436; Pubmed=8397415;
RA McFarlane I.D., Hudman D., Notacker H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two
RT inhibitory neuropeptides, Antho-Kamide and Antho-Riamide."
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC -1- FUNCTION: Inhibits spontaneous contractions in several muscle
CC groups. May be involved in the expansion phase of feeding
CC behaviour in sea anemones.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Neuron-specific.
DR Interpro: IPR001023; Hsp70.
KM Neuropeptide; Amidation.
FT CHAIN 1 4 ANTHO-RIAMIDE I.
FT CHAIN 2 4 ANTHO-RIAMIDE II.
FT MOD_RES 1 1 L-3-PHENYLACTYL.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA: 598 MW: 60441B59A0000000 CRC64:

Query Match 21.1%; Score 4; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 I 4
Db 4 I 4

RESULT 11
FAR3_HIRME STANDARD; PRT; 4 AA.
AC P42562;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide YLRF-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;
OC Aynchobellida; Hirudiniiformes; Hirudinae; Hirudo.
OX NCBI_Taxid=6421;
RN [1]
RP SEQUENCE.
RX MEDLINE=92195954; Pubmed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Galabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech."
RL Peptides 12:897-908(1991).
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KM Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
FT MOD_RES 4 4
SQ SEQUENCE 4 AA: 598 MW: 69D4073B30000000 CRC64:

Query Match 10.5%; Score 2; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 1.1e+05;
Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 I 4

```

DB 2 L 2

RESULT 12

FLRF_HIRME STANDARD; PRT; 4 AA.
AC P42561;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FLRFamide.
OS Hirudo medicinalis (Medicinal leech), and
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;
OC Arynchobdellida; Hirudiniiformes; Hirudindae; Hirudo.
OX NCBI_TaxID=6421, 27815;
RN [1]
RP SEQUENCE.
RC SPECIES-H. medicinalis;
RX MEDLINE-92195954; PubMed-1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
RN [2]
RP SEQUENCE.
RC SPECIES-H. trivolvis; TISSUE-Kidney;
RX MEDLINE-94286417; PubMed-7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma trivolvis";
RL Peptides 15:31-36(1994).
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY;
KW Neuropeptide; Amidation.
FT MOD_RES 4
SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

Query Match 10.5%; Score 2; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 1.1e+05;
Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 I 4
DB 2 L 2

RESULT 13

FLRN_ANTEL STANDARD; PRT; 4 AA.
AC P58707;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Antho-RNamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;
OC Nymphaeae; Actinidae; Anthopleura.
OX NCBI_TaxID=61110;
RN [1]
RP SEQUENCE, AND MASS-SPECTROMETRY.
RX MEDLINE-90319122; PubMed-1973541;
RA Gimmelikhuisen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
RA Reinscheid R.K., Notacker H.-P., Staley A.L.;
RT "Isolation of L-3-phenylalanyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea anemone neuropeptide containing an unusual amino-terminal blocking group";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Neuron-specific.
CC -1- MASS SPECTROMETRY: MW=549.3; METHOD-FAB.
KW Neuropeptide; Amidation.
FT MOD_RES 1
SQ SEQUENCE 1 L-3-PHENYLALANYL.

FT MOD_RES 4
SQ SEQUENCE 4 AA; 549 MW; 64540729A0000000 CRC64;

Query Match 10.5%; Score 2; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 1.1e+05;
Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 I 4
DB 2 L 2

RESULT 14

FAR4_HIRME STANDARD; PRT; 4 AA.
AC P42563;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FMRFamide-like neuropeptide YMRP-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;
OC Arynchobdellida; Hirudiniiformes; Hirudindae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.
RC SPECIES-H. medicinalis;
RX MEDLINE-92195954; PubMed-1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech";
RL Peptides 12:897-908(1991).
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY;
KW Neuropeptide; Amidation.
FT MOD_RES 4
SQ SEQUENCE 4 AA; 616 MW; 69D406B300000000 CRC64;

Query Match 5.3%; Score 1; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 1.1e+05;
Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 V 2
DB 2 M 2

RESULT 15

FMRF_MACNI STANDARD; PRT; 4 AA.
AC P01162;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE FMRFamide (Peak C) (Cardioexcitatory neuropeptide).
OS Macrocaltista nimbose (Sun-ray clam),
OS Neris virens (Sandworm),
OS Hirudo medicinalis (Medicinal leech), and
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroida;
OC Veneroidea; Veneridae; Macrocaltista.
OX NCBI_TaxID=6594, 6353, 6421, 27815;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC SPECIES-M. nimbose; TISSUE-Cerebral pedal, and visceral ganglion;
RX MEDLINE-77215956; PubMed-877582;
RA Price D.A., Greenberg M.J.;
RT "Structure of a molluscan cardioexcitatory neuropeptide";
RL Science 197:670-671(1977).
RN [2]
RP SEQUENCE, AND CHARACTERIZATION.
RC SPECIES-M. nimbose; TISSUE-ganglion;
RX MEDLINE-78012038; PubMed-909875;
RA Price D.A., Greenberg M.J.;
RT "Purification and characterization of a cardioexcitatory neuropeptide

```

RT from the central ganglia of a bivalve mollusc.";
RL Prep. Biochem. 7:261-281(1977).
RN [3]
RP SEQUENCE.
RC SPECIES=N. virens;
RX MEDLINE=90259866; PubMed=2342992;
RA Kirjula K.G., Price D.A.;
RT "Authentic FMRFamide is present in the polychaete Nereis virens.";
RL Peptides 11:75-77(1990).
RN [4]
RP SEQUENCE.
RC SPECIES=H. medicinalis;
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamde neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
RN [5]
RP SEQUENCE.
RC SPECIES=H. trivolvus; TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
trivolvus.";
RL Peptides 15:31-36(1994).
CC -I- FUNCTION: MYOACTIVE; CARDIOEXCITATORY SUBSTANCE. PHARMACOLOGICAL
CC ACTIVITIES INCLUDE AUGMENTATION, INDUCTION, AND REGULARIZATION OF
CC CARDIAC CONTRACTION.
CC -I- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR: A01426; ECKN.
DR PIR: A60418; A60418.
KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;

Query Match 5.3%; Score 1; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 1.le+05;
Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 V 2
DB 2 M 2

```

Search completed: February 16, 2003, 09:20:15
 Job time : 11 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:17:18 ; Search time 30 Seconds

(without alignments)
27.473 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhbc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 4 | 21.1 | 4 | 11 Q08433 | 008433 rattus norv |

ALIGNMENTS

RESULT 1
Q08433
ID Q08433 PRELIMINARY; PRT; 4 AA.
AC Q08433;
DT 01-NOV-1996 (TRENBLER. 01, Created)
DT 01-NOV-1996 (TRENBLER. 01, Last sequence update)
DT 01-JAN-1999 (TRENBLER. 09, Last annotation update)
DE UDP-glucuronosyltransferase, microsomal (EC 2.4.1.17) (UDPGR)
DE (Fragment).
OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91282758; PubMed=1840486;
RA SATO H., AONO S., KASHIWAMATA S., KOTAI O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT hyperbilirubinemic Gunn rat.";
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
CC -1- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC ENDOGENOUS COMPOUNDS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC BETA-D-GLUCURONOSIDE.
CC -1- SUBCELLULAR LOCATION: MICROSOME.
DR EMBL; S38636; AAB19259.1; -;
KW Transferase; Glycosyltransferase; Microsome; Multigene family.
FT NON_TER 1 1
FT NON_TER 4 4
SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 21.1%; Score 4; DB 11; Length 4;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 V 2
DB 2 V 2

Search completed: February 16, 2003, 09:20:57
Job time : 31 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 07:52:22 : Search time 80 Seconds
(without alignments)
6.663 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 19 | 100.0 | 4 | 23 | ABB76220 |
| 2 | 19 | 100.0 | 4 | 23 | AAU97413 |
| 3 | 19 | 100.0 | 4 | 23 | AAU78441 |
| 4 | 19 | 100.0 | 4 | 23 | AAU78484 |
| 5 | 19 | 100.0 | 5 | 23 | ABB76215 |
| 6 | 19 | 100.0 | 5 | 23 | AAU78433 |
| 7 | 19 | 100.0 | 5 | 23 | AAU78485 |
| 8 | 19 | 100.0 | 6 | 23 | ABB76214 |
| 9 | 19 | 100.0 | 6 | 23 | AAU78486 |
| 10 | 19 | 100.0 | 7 | 23 | ABB76213 |

| | | | | | | |
|----|----|-------|----|----|----------|--------------------|
| 11 | 19 | 100.0 | 7 | 23 | AAU78434 | Inhibitor of apopt |
| 12 | 19 | 100.0 | 7 | 23 | AAU78487 | Smac-7 AV peptid. |
| 13 | 19 | 100.0 | 8 | 22 | ABP18027 | HIV B58 super moti |
| 14 | 19 | 100.0 | 8 | 22 | ABB76212 | Human smac (DIABLO |
| 15 | 19 | 100.0 | 9 | 22 | AAU49855 | Human endostatin p |
| 16 | 19 | 100.0 | 9 | 23 | ABB76209 | Human smac (DIABLO |
| 17 | 19 | 100.0 | 9 | 23 | ABB76210 | Human smac (DIABLO |
| 18 | 19 | 100.0 | 9 | 23 | ABB76211 | Human smac (DIABLO |
| 19 | 19 | 100.0 | 9 | 23 | ABB76216 | Human smac (DIABLO |
| 20 | 19 | 100.0 | 9 | 23 | ABB76225 | Human smac (DIABLO |
| 21 | 19 | 100.0 | 9 | 23 | ABB76229 | Human smac (DIABLO |
| 22 | 19 | 100.0 | 9 | 23 | AAU78440 | Smac-caspase 8-cle |
| 23 | 19 | 100.0 | 10 | 22 | ABP18029 | HIV B58 super moti |
| 24 | 19 | 100.0 | 10 | 23 | ABB76228 | Fluorocelinated sm |
| 25 | 19 | 100.0 | 12 | 21 | AAU98122 | Improved peptide u |
| 26 | 19 | 100.0 | 15 | 19 | AAU46459 | Site of the fusion |
| 27 | 19 | 100.0 | 15 | 21 | AAU84176 | Amino acid sequenc |
| 28 | 19 | 100.0 | 16 | 20 | AAU96271 | B-cell epitope of |
| 29 | 19 | 100.0 | 20 | 21 | AAU78398 | Human papillomavir |
| 30 | 19 | 100.0 | 20 | 23 | ABB76208 | Human smac (DIABLO |
| 31 | 19 | 100.0 | 22 | 22 | ABA49869 | Human endostatin p |
| 32 | 19 | 100.0 | 23 | 20 | AAU40072 | Peptide sequence d |
| 33 | 19 | 100.0 | 29 | 23 | AAU91218 | Human E1-E2 ATPase |
| 34 | 19 | 100.0 | 30 | 23 | AAU78435 | Inhibitor of apopt |
| 35 | 19 | 100.0 | 31 | 22 | ABG09310 | Novel human diagno |
| 36 | 19 | 100.0 | 35 | 22 | AAU10938 | Mouse massl protel |
| 37 | 19 | 100.0 | 35 | 23 | AAU78439 | Inhibitor of apopt |
| 38 | 19 | 100.0 | 36 | 22 | ABG09529 | Novel human diagno |
| 39 | 19 | 100.0 | 39 | 23 | AAU78436 | Inhibitor of apopt |
| 40 | 19 | 100.0 | 40 | 23 | AAU78430 | Inhibitor of apopt |
| 41 | 19 | 100.0 | 46 | 21 | AAU56278 | Arabidopsis thalia |
| 42 | 19 | 100.0 | 46 | 21 | AAU56298 | Arabidopsis thalia |
| 43 | 19 | 100.0 | 46 | 21 | AAU61228 | Arabidopsis thalia |
| 44 | 19 | 100.0 | 46 | 21 | AAU61231 | Arabidopsis thalia |
| 45 | 19 | 100.0 | 46 | 22 | AAU14206 | Human novel protel |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| ABB76220 | |
| ID | ABB76220 standard; Peptide; 4 AA. |
| XX | |
| AC | ABB76220; |
| XX | |
| DT | 09-AUG-2002 (first entry) |
| XX | |
| DE | Human smac (DIABLO) derived peptide. |
| XX | |
| KW | DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis; |
| KW | human; cancer; cytosolic; mutant; mutlein. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FH | Key |
| FT | Modified-site 4 Location/Qualifiers |
| XX | |
| PN | WO200230959-A2. |
| XX | |
| PD | 18-APR-2002. |
| XX | |
| PF | 12-OCT-2001; 2001WO-US32121. |
| XX | |
| PR | 13-OCT-2000; 2000US-0687549. |
| XX | |
| PA | (ABBO) ABBOTT LAB. |
| XX | |
| PI | Fesik SW, Meadows RP, Betz SP, Liu Z, Olejniczak ET, Sun C; |
| XX | |
| DR | WPI; 2002-444169/47. |
| XX | |

PT Novel peptide derived from wild-type human second mitochondria derived
 PT activator of caspase protein useful for identifying candidate
 PT substances to kill cancerous cells -
 XX
 PS Example 1; Page 15; 26pp; English.
 XX
 CC The present sequence is a peptide derived from wild-type human
 CC second mitochondria derived activator of caspase (smac), also known
 CC as direct inhibitor of apoptosis binding protein with low PI
 CC (DIABLO). Claimed smac (DIABLO)-derived peptides (see AB876208-19)
 CC bind to the BIR2 and BIR3 domain of XIAP, an inhibitor of apoptosis
 CC protein (IAP) family member. They can be used to identify
 CC candidate substances which induce or promote apoptosis in cells.
 CC The assay involves determination of the ability of candidate
 CC compounds to disrupt the binding interaction between a smac
 CC (DIABLO) peptide and an IAP family member.
 XX
 SQ Sequence 4 AA;
 XX
 Query Match 100.0%; Score 19; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AVPI 4
 Db 1 AVPI 4
 XX
 RESULT 2
 AAU97413
 ID AAU97413 standard; Peptide; 4 AA.
 XX
 AC AAU97413;
 XX
 DT 30-JUL-2002 (first entry)
 XX
 DE Human Smac (second mitochondria-derived activator of caspases) peptide.
 XX
 KW Human; synthetic tetrapeptide; mimetic; inhibitor of apoptosis; IAP;
 KW baculoviral IAP repeat; BIR: cell proliferative disease; cancer;
 KW Alzheimer's disease; second mitochondria-derived activator of caspases;
 KW stroke; arthritis; Smac.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 PN WO20022675-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 28-SEP-2001; 2001WO-US30567.
 XX
 PR 29-SEP-2000; 2000US-236574P.
 XX
 PR 20-DEC-2000; 2000US-256830P.
 XX
 PA (UYPR-) UNIV PRINCETON.
 XX
 PI Shi Y;
 XX
 DR WPI; 2002-416473/44.
 XX
 PT Novel synthetic tetrapeptide, its mimetic or a compound useful for
 PT rational drug design and in stimulating apoptosis in cell, binds an
 PT inhibitor of apoptosis protein and relieves inhibition of caspase -
 XX
 PS Claim 6; Page 43; 62pp; English.
 XX
 CC The present invention relates to a new synthetic tetrapeptide, its
 CC non-peptide or partial peptide mimetic or a compound that binds an
 CC inhibitor of apoptosis protein (IAP) and relieves IAP-mediated inhibition
 CC of caspase activity, where the tetrapeptide binds a surface groove within
 CC a baculoviral IAP repeat (BIR) domain of the IAP. The invention can be
 CC used for stimulating apoptosis in a cultured cell and for screening for a

CC compound that binds an IAP at a surface groove within a BIR domain.
 CC The synthetic tetrapeptide is a partial peptide or non-peptide mimetic
 CC and is useful in making a drug suitable for treating cell proliferative
 CC disease especially cancer in a mammal by promoting apoptosis in
 CC proliferatively diseased cells. Agonists identified by the invention are
 CC useful in treating cell proliferative diseases like cancer and
 CC antagonists identified by the invention are useful in treating
 CC Alzheimer's disease, stroke and arthritis. The present amino acid
 CC sequence represents the human Smac (second mitochondria-derived
 CC activator of caspases) peptide. This sequence is a synthetic
 CC tetrapeptide of the invention.
 XX
 SQ Sequence 4 AA;
 XX
 Query Match 100.0%; Score 19; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AVPI 4
 Db 1 AVPI 4
 XX
 RESULT 3
 AAU78441
 ID AAU78441 standard; Peptide; 4 AA.
 XX
 AC AAU78441;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Inhibitor of apoptosis (IAP) protein Smac, peptide #1.
 XX
 DE Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
 KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
 KW neoplastic cell; tumour.
 XX
 OS Homo sapiens.
 OS
 PN WO200216418-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 24-AUG-2001; 2001WO-US26492.
 XX
 PR 24-AUG-2000; 2000US-227735P.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Alnemri ES;
 XX
 DR WPI; 2002-304115/34.
 XX
 PT Novel Smac peptides and polynucleotides encoding the peptides, useful
 PT for stimulating apoptosis in neoplastic or tumour cell which
 PT overexpresses inhibitor of caspase, and for identifying apoptosis
 PT modulating compounds -
 XX
 PS Claim 8; Page 12; 78pp; English.
 XX
 CC The invention relates to an isolated Smac peptide or polypeptide (1)
 CC and an isolated nucleic acid (II) encoding (I). Also described is a
 CC method of identifying a compound that inhibits apoptosis, comprising:
 CC (a) separately contacting several cell populations expressing a
 CC cytosolic Smac (a Smac isoform that begins with MKSDPYF sequence,
 CC replacing the mitochondrial targeting sequence (residues 1-55 of (1)),
 CC and residues 56-60 of (1)) and an inhibitor of BID (Bcl2 interacting
 CC domain) with a compound to be tested for apoptotic inhibiting activity;
 CC (b) incubating the cell populations with a direct stimulus of the cell
 CC death pathway; and (c) measuring the specific apoptotic activity of the
 CC cell populations, where inhibition of the specific apoptotic activity is
 CC indicative that the compound is an inhibitor of apoptosis. (1) and (II)
 CC are useful for inducing apoptosis in a cell. The Smac polypeptide and

CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
 CC tumor cell which overexpresses an inhibitor of caspase, where the
 CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
 CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
 CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
 CC mediated apoptosis which involves contacting a cell transformed or
 CC transfected with a vector expressing (I) with a candidate inhibitor or
 CC candidate enhancer; and detecting cell viability, where an increase in
 CC cell viability indicates the presence of an inhibitor and a decrease in
 CC cell viability indicates the presence of an enhancer. Optionally, the
 CC method involves detecting the presence of large and small caspase
 CC subunits after contacting cell transformed with the vector expressing
 CC (I), with the candidate compound. A decrease in processing indicates the
 CC presence of an inhibitor and an increase in the processing indicates the
 CC presence of an enhancer. Preferably, the large and small subunits of
 CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
 CC identifying a compound that inhibits Smac binding to Smac-binding
 CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
 CC or a full-length IAP). (II) is useful in gene therapy techniques. The
 CC present sequence represents the amino acid sequence of Smac peptide #1.
 CC
 SQ Sequence 4 AA;

Query Match 100.0%; Score 19; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
 ||||
 Db 1 AVPI 4

RESULT 4
 AAU78484
 ID AAU78484 standard; Peptide; 4 AA.

AC AAU78484;

DT 18-JUN-2002 (first entry)

DE Smac-4 AV peptoid.

KM Apoptosis; cytosolic; apoptotic; AV peptoid; melanoma; lymphoma;
 KM Inhibitor of Apoptosis protein; IAP; procaspase-3; tumour cell;

KM breast cancer; prostate cancer; lung cancer; pancreatic cancer; smac-4;
 KM gastric cancer; colon cancer; ovarian cancer; renal cancer; hepatoma;

KM sarcoma; smac; second mitochondria-derived activator of caspases.

OS Synthetic.

PN WO200216402-A2.

PD 28-FEB-2002.

PF 23-AUG-2001; 2001WO-US41869.

PR 23-AUG-2000; 2000US-0645075.

PA (TEXA) UNIV TEXAS SYSTEM.

PI Wang X, Du C;

DR WPI; 2002-280909/32.

PT Composition for enhancing the apoptosis of pathogenic cells,
 PT particularly tumour cells, e.g. breast cancer, prostate cancer, lung

PT cancer, colon cancer, ovarian cancer or sarcoma, comprises apoptotic
 PT compounds -

XX Example 9; Page 28; 40pp; English.

CC This invention relates to a method for induction of apoptosis in
 CC pathogenic cells. The method comprises a novel pharmaceutical

CC composition which comprises an AV peptoid in dosage form and a
 CC pharmaceutical carrier, where the AV peptoid comprises a peptide that
 CC interacts with or inhibits the activity of an inhibitor of Apoptosis
 CC protein (IAP) as measured by IAP binding, procaspase-3 activation or
 CC promotion of apoptosis. The peptoids of the invention are used to
 CC inhibit an inhibitor of Apoptosis protein (IAP). Compositions containing
 CC these peptides are useful for enhancing the apoptosis of pathogenic
 CC cells, particularly tumour cells, e.g. breast cancer, prostate cancer,
 CC lung cancer, pancreatic cancer, gastric cancer, colon cancer, ovarian
 CC cancer, renal cancer, hepatoma, melanoma, lymphoma or sarcoma. The
 CC composition is particularly useful for promoting cell death. The
 CC present sequence represents an AV peptoid (smac-4) used to inhibit
 CC second mitochondria-derived activator of caspases (smac) using the
 CC method of the invention. Smac interacts with and eliminates the activity
 CC of a number of IAP's and as such inhibiting its activity allows the
 CC induction of apoptosis.

Query Match 100.0%; Score 19; DB 23; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
 ||||
 Db 1 AVPI 4

RESULT 5
 ABB76215
 ID ABB76215 standard; Peptide; 5 AA.

AC ABB76215;

DT 09-AUG-2002 (first entry)

DE Human smac (DIABLO) derived peptide.

KM DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
 KM human; cancer; cytosolic; mutant; mutleln.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Modified-site 5 /note="Optional C-terminal protecting group,
 FT e.g. C-terminal amide"

FT WO200230959-A2.

PN 18-APR-2002.

PD 12-OCT-2001; 2001WO-US32121.

PF 13-OCT-2000; 2000US-0687549.

PR (ABBO) ABBOTT LAB.

PA Fesik SM, Meadows RP, Betz SP, Liu Z, Olejniczak ET, Sun C;

PI WPI; 2002-444169/47.

DR Novel peptide derived from wild-type human second mitochondria derived

PT activator of caspase protein useful for identifying candidate
 PT substances to kill cancerous cells -

PS Claim 5; Page 7; 26pp; English.

XX The present sequence is a peptide derived from wild-type human
 XX second mitochondria derived activator of caspase (smac), also known

CC as direct inhibitor of apoptosis binding protein with low PI
 CC (DIABLO). The peptide is one of 12 claimed smac (DIABLO)-derived

CC peptides (see ABB76208-19) which bind to the Bir2 and Bir3 domain

CC of XIAP, an inhibitor of apoptosis protein (IAP) family member.
 CC Kd values for Bir-3 and Bir-2 are 0.64 +/- 0.07 uM and 5.5 +/- 0.5
 CC uM, respectively, for the present (C-terminally amidated) peptide,
 CC compared with 0.42 +/- 0.02 uM and 2.3 +/- 0.3 uM, respectively,
 CC for full-length smac, showing that a peptide of 5 amino acids
 CC retained full binding affinity to the Bir3 domain of XIAP.
 CC Modification of the N-terminal alanine destroyed binding affinity
 CC to XIAP, and mutation of the valine, proline or isoleucine also
 CC caused some loss of binding. The claimed peptides can be used to
 CC identify candidate substances which induce or promote apoptosis in
 CC cells. The assay involves determination of the ability of
 CC candidate compounds to disrupt the binding interaction between a
 CC smac (DIBAL0) peptide and an IAP family member.

CC Sequence 5 AA:

Query Match 100.0%; Score 19; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
 ||||
 DB 1 AVPI 4

RESULT 6

AAU78433
 ID AAU78433 standard; Peptide; 5 AA.

AC AAU78433;

DT 18-JUN-2002 (first entry)

DE Inhibitor of apoptosis (IAP) protein Smac, long isoform peptide.

XX Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
 KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
 KW neoplastic cell; tumour.

OS Homo sapiens.

XX WO200216418-A2.

PN 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US26492.

XX 24-AUG-2000; 2000US-227735P.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Alnemrl ES;

XX WPI: 2002-304115/34.

XX Novel Smac peptides and polynucleotides encoding the peptides, useful
 PT for stimulating apoptosis in neoplastic or tumour cell which
 PT overexpresses inhibitor of caspase, and for identifying apoptosis
 PT modulating compounds -

XX Example 5; Fig 12; 78pp; English.

XX The invention relates to an isolated smac peptide or polypeptide (I)
 CC and an isolated nucleic acid (II) encoding (I). Also described is a
 CC method of identifying a compound that inhibits apoptosis, comprising:
 CC (a) separately contacting several cell populations expressing a
 CC cytosolic smac (a smac isoform that begins with MRSDFV sequence,
 CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
 CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
 CC domain) with a compound to be tested for apoptotic inhibiting activity;
 CC (b) incubating the cell populations with a direct stimulus of the cell
 CC death pathway; and (c) measuring the specific apoptotic activity of the
 CC cell populations, where inhibition of the specific apoptotic activity is

CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
 CC are useful for inducing apoptosis in a cell. The smac polypeptide and
 CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
 CC tumour cell which overexpresses an inhibitor of caspase, where the
 CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
 CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
 CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
 CC mediated apoptosis which involves contacting a cell transformed or
 CC transfected with a vector expressing (I) with a candidate inhibitor or
 CC candidate enhancer; and detecting cell viability, where an increase in
 CC cell viability indicates the presence of an inhibitor and a decrease in
 CC cell viability indicates the presence of an enhancer. Optionally, the
 CC method involves detecting the presence of large and small caspase
 CC subunits after contacting cell transformed with the vector expressing
 CC (I), with the candidate compound. A decrease in processing indicates the
 CC presence of an inhibitor and an increase in the processing indicates the
 CC presence of an enhancer. Preferably, the large and small subunits of
 CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
 CC identifying a compound that inhibits smac binding to smac-binding
 CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
 CC or a full-length IAP). (II) is useful in gene therapy techniques. The
 CC present sequence represents the amino acid sequence of smac protein long
 CC isoform peptide.

CC Sequence 5 AA:

Query Match 100.0%; Score 19; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
 ||||
 DB 1 AVPI 4

RESULT 7

AAU78485
 ID AAU78485 standard; Peptide; 5 AA.

AC AAU78485;

DT 18-JUN-2002 (first entry)

DE Smac-5 AV peptoid.

XX Apoptosis; cytosolic; AV peptoid; melanoma; lymphoma;

XX Inhibitor of Apoptosis protein; IAP; procaspase-3; tumour cell;

XX breast cancer; prostate cancer; lung cancer; pancreatic cancer; smac-5;

XX gastric cancer; colon cancer; ovarian cancer; renal cancer; hepatoma;

XX sarcoma; smac; second mitochondria-derived activator of caspases.

XX Synthetic.

XX WO200216402-A2.

XX 28-FEB-2002.

XX 23-AUG-2001; 2001WO-US41869.

XX 23-AUG-2000; 2000US-0645075.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Wang X, Du C;

XX WPI: 2002-280909/32.

XX Composition for enhancing the apoptosis of pathogenic cells,
 PT particularly tumour cells, e.g. breast cancer, prostate cancer, lung
 PT cancer, colon cancer, ovarian cancer or sarcoma, comprises apoptotic
 PT compounds -
 XX Example 9; Page 28; 40pp; English.

XX This invention relates to a method for induction of apoptosis in
CC pathogenic cells. The method comprises a novel pharmaceutical
CC composition which comprises an AV peptide in dosage form and a
CC pharmaceutical carrier, where the AV peptide comprises a peptide that
CC interacts with or inhibits the activity of an inhibitor of Apoptosis
CC protein (IAP) as measured by IAP binding, procaspase-3 activation or
CC promotion of apoptosis. The peptides of the invention are used to
CC inhibit an inhibitor of Apoptosis protein (IAP). Compositions containing
CC these peptides are useful for enhancing the apoptosis of pathogenic
CC cells, particularly tumour cells, e.g. breast cancer, prostate cancer,
CC lung cancer, pancreatic cancer, gastric cancer, colon cancer, ovarian
CC cancer, renal cancer, hepatoma, melanoma, lymphoma or sarcoma. The
CC composition is particularly useful for promoting cell death. The
CC present sequence represents an AV peptide (smac-5) used to inhibit
CC second mitochondria-derived activator of caspases (smac) using the
CC method of the invention. Smac interacts with and eliminates the activity
CC of a number of IAP's and as such inhibiting its activity allows the
CC induction of apoptosis.

SQ Sequence 5 AA:
Query Match 100.0%; Score 19; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 8
ABB76214
ID ABB76214 standard; Peptide: 6 AA.
XX
AC ABB76214;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human smac (DIABLO) derived peptide.
XX
KW DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
KW human; cancer; cytostatic; mutant; mutlein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 6 /note="optional C-terminal protecting group,
FT e.g. C-terminal amide"
XX
PN WO200230959-A2.
XX
PD 18-APR-2002.
XX
PE 12-OCT-2001; 2001WO-US32121.
XX
PR 13-OCT-2000; 2000US-0687549.
XX
PA (ABBO) ABBOT LAB.
XX
PI Fesik SM, Meadows RP, Betz SP, Liu Z, Olejniczak ET, Sun C;
XX WPI; 2002-444169/47.
XX
PT Novel peptide derived from wild-type human second mitochondria derived
XX activator of caspase protein useful for identifying candidate
XX substances to kill cancerous cells -
PS Claim 5; Page 7; 26pp; English.
XX
CC The present sequence is a peptide derived from wild-type human
CC second mitochondria derived activator of caspase (smac), also known

CC as direct inhibitor of apoptosis binding protein with low pI
CC (DIABLO). The peptide is one of 12 claimed smac (DIABLO)-derived
CC peptides (see ABB76208-19) which bind to the Bir2 and Bir3 domain
CC of XIAP, an inhibitor of apoptosis protein (IAP) family member.
CC Kd values for Bir-3 and Bir-2 are 0.80 +/- 0.2 uM and 8.9 +/- 0.6
CC uM, respectively, for the present C-terminally amidated peptide,
CC compared with 0.42 +/- 0.02 uM and 2.3 +/- 0.3 uM, respectively,
CC for full-length smac. Modification of the N-terminal alanine
CC destroys binding affinity to XIAP, and mutation of the valine,
CC proline or isoleucine also causes some loss of binding. Amino
CC acids C-terminal to the isoleucine are not important for binding.
CC The claimed peptides can be used to identify candidate substances
CC which induce or promote apoptosis in cells. The assay involves
CC determination of the ability of candidate compounds to disrupt the
CC binding interaction between a smac (DIABLO) peptide and an IAP
CC family member.

SQ Sequence 6 AA:
Query Match 100.0%; Score 19; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 9
AAU78486
ID AAU78486 standard; Peptide: 6 AA.
XX
AC AAU78486;
XX
DT 18-JUN-2002 (first entry)
XX
DE Smac-6 AV peptide.
XX
KW Apoptosis; cytostatic; apoptotic; AV peptide; melanoma; lymphoma;
KW Inhibitor of Apoptosis protein; IAP; procaspase-3; tumour cell;
KW breast cancer; prostate cancer; lung cancer; pancreatic cancer; smac-6;
KW gastric cancer; colon cancer; ovarian cancer; renal cancer; hepatoma;
KW sarcoma; smac; second mitochondria-derived activator of caspases.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 6 /note="optional C-terminal protecting group,
FT e.g. C-terminal amide"
XX
PN WO200216402-A2.
XX
PD 28-FEB-2002.
XX
PE 23-AUG-2001; 2001WO-US41869.
XX
PR 23-AUG-2000; 2000US-0645075.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Wang X, Du C;
XX WPI; 2002-280909/32.
XX
PT Composition for enhancing the apoptosis of pathogenic cells,
PT particularly tumour cells, e.g. breast cancer, prostate cancer, lung
PT cancer, colon cancer, ovarian cancer or sarcoma, comprises apoptotic
PT compounds -
XX
PS Example 9; Page 28; 40pp; English.
XX
CC This invention relates to a method for induction of apoptosis in
CC pathogenic cells. The method comprises a novel pharmaceutical
CC composition which comprises an AV peptide in dosage form and a
CC pharmaceutical carrier, where the AV peptide comprises a peptide that
CC interacts with or inhibits the activity of an inhibitor of Apoptosis
CC protein (IAP) as measured by IAP binding, procaspase-3 activation or

CC promotion of apoptosis. The peptidoids of the invention are used to
 CC inhibit an inhibitor of Apoptosis protein (IAP). Compositions containing
 CC these peptidoids are useful for enhancing the apoptosis of pathogenic
 CC cells, particularly tumour cells, e.g. breast cancer, prostate cancer,
 CC lung cancer, pancreatic cancer, gastric cancer, colon cancer, ovarian
 CC cancer, renal cancer, hepatoma, melanoma, lymphoma or sarcoma. The
 CC composition is particularly useful for promoting cell death. The
 CC present sequence represents an AV peptid (smac-6) used to inhibit
 CC second mitochondria-derived activator of caspases (smac) using the
 CC method of the invention. Smac interacts with and eliminates the activity
 CC of a number of IAP's and as such inhibiting its activity allows the
 CC induction of apoptosis.

CC Sequence 6 AA:

Query Match 100.0%; Score 19; DB 23; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
 ||||
 Db 1 AVPI 4

RESULT 10

AB76213
 ID ABB76213 standard; Peptide: 7 AA.

AC ABB76213;

DT 09-AUG-2002 (first entry)

DE Human smac (DIABLO) derived peptide.

KW DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
 KM human; cancer; cytostatic; mutant; mutein.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Modified-site 7 /note="optional C-terminal protecting group,
 e.g. C-terminal amide"

PN WO200230959-A2.

PD 18-APR-2002.

PF 12-OCT-2001; 2001WO-US32121.

PR 13-OCT-2000; 2000US-0687549.

PA (ABBO) ABBOTT LAB.

PI Fesik SM, Meadows RP, Betz SP, Liu Z, Olejniczak ET, Sun C;

DR MPI; 2002-444169/47.

PT Novel peptide derived from wild-type human second mitochondria derived
 PT activator of caspase protein useful for identifying candidate
 PT substances to kill cancerous cells -

XX Claim 5; Page 7; 26pp; English.

CC The present sequence is a peptide derived from wild-type human
 CC second mitochondria derived activator of caspase (smac), also known
 CC as direct inhibitor of apoptosis binding protein with low pI
 CC (DIABLO). The peptide is one of 12 claimed smac (DIABLO)-derived
 CC peptides (see ABB76208-19) which bind to the Bir2 and Bir3 domain
 CC of XIAP, an inhibitor of apoptosis protein (IAP) family member.
 CC Kd values for Bir-3 and Bir-2 are 0.70 +/- 0.09 uM and 9.4 +/- 0.6
 CC uM, respectively, for the present (C-terminally amidated) peptide,
 CC compared with 0.42 +/- 0.02 uM and 2.3 +/- 0.3 uM, respectively,

CC for full-length smac. Modification of the N-terminal alanine
 CC destroys binding affinity to XIAP, and mutation of the valine,
 CC proline or isoleucine also causes some loss of binding. Amino
 CC acids C-terminal to the isoleucine are not important for binding.
 CC The claimed peptides can be used to identify candidate substances
 CC which induce or promote apoptosis in cells. The assay involves
 CC determination of the ability of candidate compounds to disrupt the
 CC binding interaction between a smac (DIABLO) peptide and an IAP
 CC family member.

CC Sequence 7 AA:

Query Match 100.0%; Score 19; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
 ||||
 Db 1 AVPI 4

RESULT 11

AAU78434
 ID AAU78434 standard; Peptide: 7 AA.

AC AAU78434;

DT 18-JUN-2002 (first entry)

DE Inhibitor of apoptosis (IAP) protein Smac, mutant Smac-N7.

KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
 KM Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;

KW neoplastic cell; mutant; tumour.

OS Homo sapiens.

OS Synthetic.

PN WO200216418-A2.

PD 28-FEB-2002.

PF 24-AUG-2001; 2001WO-US26492.

PR 24-AUG-2000; 2000US-227735P.

PA (UYJE-) UNIV JEFFERSON THOMAS.

PI Alnemri ES;

DR MPI; 2002-304115/34.

PT Novel Smac peptides and polynucleotides encoding the peptides, useful
 PT for stimulating apoptosis in neoplastic or tumour cell which
 PT overexpresses inhibitor of caspase, and for identifying apoptosis
 PT modulating compounds -

XX Example 3; Fig 7; 78pp; English.

CC The invention relates to an isolated Smac peptide or polypeptide (I)
 CC and an isolated nucleic acid (II) encoding (I). Also described is a
 CC method of identifying a compound that inhibits apoptosis, comprising:
 CC (a) separately contacting several cell populations expressing a
 CC cytosolic Smac (a Smac isoform that begins with MKSDPYF sequence,
 CC replacing the mitochondrial targeting sequence (residues 1-53 of (I)),
 CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
 CC domain) with a compound to be tested for apoptotic inhibiting activity;
 CC (b) incubating the cell populations with a direct stimulus of the cell
 CC death pathway; and (c) measuring the specific apoptotic activity of the
 CC cell populations, where inhibition of the specific apoptotic activity is
 CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
 CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
 CC polynucleotide are useful for stimulating apoptosis in a neoplastic or

CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (I) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac mutant
CC Smac-N7.
CC
XX

SQ Sequence 7 AA;

Query Match

Best Local Similarity 100.0%; Score 19; DB 23; Length 7;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4

DB 1 AVPI 4

RESULT 12

AAU78487

ID AAU78487 standard; Peptide; 7 AA.

AC AAU78487;

DT 18-JUN-2002 (first entry)

DE Smac-7 AV peptoid.

XX Apoptosis; cytosolic; apoptotic; AV peptoid; melanoma; lymphoma;
KW Inhibitor of Apoptosis protein; IAP; procaspase-3; tumour cell;
KW breast cancer; prostate cancer; lung cancer; pancreatic cancer; smac-7;
KW gastric cancer; colon cancer; ovarian cancer; renal cancer; hepatoma;
KW sarcoma; smac; second mitochondria-derived activator of caspases.
XX

OS Synthetic.

PN WO200216402-A2.

PD 28-FEB-2002.

PF 23-AUG-2001; 2001WO-US41869.

PR 23-AUG-2000; 2000US-0645075.

PA (TEXA) UNIV TEXAS SYSTEM.

PI Wang X, Du C;

DR WPI; 2002-280909/32.

PT Composition for enhancing the apoptosis of pathogenic cells,
PT particularly tumour cells, e.g. breast cancer, prostate cancer, lung
PT cancer, colon cancer, ovarian cancer or sarcoma, comprises apoptotic
PT compounds -
XX

PS Example 9; Page 28; 40pp; English.

XX This invention relates to a method for induction of apoptosis in
CC pathogenic cells. The method comprises a novel pharmaceutical

CC composition which comprises an AV peptoid in dosage form and a
CC pharmaceutical carrier, where the AV peptoid comprises a peptide that
CC interacts with or inhibits the activity of an inhibitor of Apoptosis
CC protein (IAP) as measured by IAP binding, procaspase-3 activation or
CC promotion of apoptosis. The peptoides of the invention are used to
CC inhibit an inhibitor of Apoptosis protein (IAP). Compositions containing
CC these peptoides are useful for enhancing the apoptosis of pathogenic
CC cells, particularly tumour cells, e.g. breast cancer, prostate cancer,
CC lung cancer, pancreatic cancer, gastric cancer, colon cancer, ovarian
CC cancer, renal cancer, hepatoma, melanoma, lymphoma or sarcoma. The
CC composition is particularly useful for promoting cell death. The
CC present sequence represents an AV peptoid (smac-7) used to inhibit
CC second mitochondria-derived activator of caspases (smac) using the
CC method of the invention. Smac interacts with and eliminates the activity
CC of a number of IAP's and as such inhibiting its activity allows the
CC induction of apoptosis.
CC
XX

SQ Sequence 7 AA;

Query Match

Best Local Similarity 100.0%; Score 19; DB 23; Length 7;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4

DB 1 AVPI 4

RESULT 13

ABP18027

ID ABP18027 standard; Peptide; 8 AA.

AC ABP18027;

DT 15-JUL-2002 (first entry)

DE HIV B58 super motif pol peptide #20.

XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr;
KW VPU; Vif; tat; cytototoxic T lymphocyte; CTL; Immune response; epitope;
KW antigen; vaccine; HIV infection; Immunisation; virucide.
XX

OS Human immunodeficiency virus type 1.

PN WO200124810-A1.

PD 12-APR-2001.

PF 05-OCT-2000; 2000WO-US27766.

PR 05-OCT-1999; 99US-0412863.

PA (EPIM-) EPIMUNE INC.

PI Sette A, Sidney J, Livingston S, Livingston BD, Chesnut R;
KW Baker DM, Cells E, Kubo RT, Grey HM;

DR WPI; 2001-354887/37.

PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1 -
XX

PS Claim 32; Page 238; 448pp; English.

XX The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABU25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I)
CC may be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,

CC allowing for immune escape due to mutations. The groups for inclusion in
 CC an group-passed vaccine may be selected from conserved regions of viral or
 CC tumour-associated antigens, which therefore reduces the likelihood of
 CC escape mutants. Furthermore, immunosuppressive groups that may be present
 CC in whole antigens can be avoided with the use of group-based vaccines.
 CC An additional advantage of an group-based vaccine approach is the ability
 CC to combine selected groups (CTL and HTL), and further, to modify the
 CC composition of the groups, achieving, for example, enhanced
 CC immunogenicity. Accordingly, the immune response can be modulated, as
 CC appropriate, for the target disease. Similar engineering of the response
 CC is not possible with traditional approaches. ABB11501 to ABP25412
 CC represent peptide sequences used in the exemplification of the present
 CC invention.

CC Sequence 8 AA;

Query Match 100.0%; Score 19; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
 ||||
 DB 5 AVPI 8

RESULT 14

ABB76212
 ID ABB76212 standard; Peptide: 8 AA.

AC ABB76212;

DT 09-AUG-2002 (first entry)

DE Human smac (DIABLO) derived peptide.

KM DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
 KM human; cancer; cytostatic; mutant; mutain.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Modified-site 8 /note= "optional C-terminal protecting group,
 e.g. C-terminal amide"

PN WO200230959-A2.

PD 18-APR-2002.

PF 12-OCT-2001; 2001WO-US3121.

PR 13-OCT-2000; 2000US-0687549.

PA (ABBO) ABBOTT LAB.

PI Fesik SM, Meadows RP, Betz SP, Liu Z, Olejniczak ET, Sun C;

PS WPI: 2002-444169/47.

PT Novel peptide derived from wild-type human second mitochondria derived
 PT activator of caspase protein useful for identifying candidate
 PT substances to kill cancerous cells -

PS Claim 5; Page 7; 26pp; English.

CC The present sequence is a peptide derived from wild-type human
 CC second mitochondria derived activator of caspase (smac), also known
 CC as direct inhibitor of apoptosis binding protein with low pI
 CC (DIABLO). The peptide is one of 12 claimed smac (DIABLO)-derived
 CC peptides (see ABB76208-19) which bind to the Bir2 and Bir3 domain
 CC of XIAP, an inhibitor of apoptosis protein (IAP) family member.
 CC Kd values for Bir-3 and Bir-2 are 0.80 +/- 0.03 uM and 13 +/- 0.3
 CC uM, respectively, for the present (C-terminally amidated) peptide,

CC compared with 0.42 +/- 0.02 uM and 2.3 +/- 0.3 uM, respectively,
 CC for full-length smac. Modification of the N-terminal alanine
 CC destroys binding affinity to XIAP, and mutation of the valine,
 CC proline or isoleucine also causes some loss of binding. Amino
 CC acids C-terminal to the isoleucine are not important for binding.
 CC The claimed peptides can be used to identify candidate substances
 CC which induce or promote apoptosis in cells. The assay involves
 CC determination of the ability of candidate compounds to disrupt the
 CC binding interaction between a smac (DIABLO) peptide and an IAP
 CC family member.

CC Sequence 8 AA;

Query Match 100.0%; Score 19; DB 23; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
 ||||
 DB 1 AVPI 4

RESULT 15

AAB49855
 ID AAB49855 standard; Peptide: 9 AA.

AC AAB49855;

DT 02-MAR-2001 (first entry)

DE Human endostatin peptide fragment SEQ ID NO: 70.

KM Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken;
 KM cancer; inflammation; angiogenesis-dependent disease.

OS Homo sapiens.

PN WO200067771-A1.

PD 16-NOV-2000.

PF 02-MAY-2000; 2000WO-US12063.

PR 06-MAY-1999; 99US-0132907.

PR 14-JUL-1999; 99US-0353333.

PA (BURN-) BURNHAM INST.

PI Vuori K;

PS WPI: 2001-040937/05.

PT Endostatin peptide comprising at least four endostatin amino acid
 PT residues are e.g. angiogenesis inhibitors for treating cancer and
 PT diabetic retinopathy -

PS Disclosure: Page 136; 146pp; English.

CC The present invention provides endostatin peptides which can be used in
 CC the modulation of angiogenesis. This is useful in the treatment of
 CC cancers, inflammation, rheumatoid arthritis, chronic articular
 CC rheumatism, psoriasis, disorders associated with inappropriate invasion of
 CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy
 CC of prematurity, macular degeneration, corneal graft rejection,
 CC retrolental fibroplasia, rubeosis, capillary proliferation in
 CC atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent
 CC diseases include Osler-Weber syndrome, myocardial angiogenesis, plaque
 CC neovascularisation, telangiectasia, haemophilic joints and wound
 CC granulation. In addition, the peptides can be used as birth control
 CC agents.

CC Sequence 9 AA;

Query Match 100.0%; Score 19; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVPI 4
 ||||
Db 2 AVPI 5

Search completed: February 16, 2003, 09:06:44
Job time : 81 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:05:18 ; Search time 38 Seconds

(without alignments)
3.097 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents-AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCtUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 19 | 100.0 | 15 | 1 | US-08-449-207-4 |
| 2 | 19 | 100.0 | 20 | 2 | US-08-850-910A-20 |
| 3 | 19 | 100.0 | 41 | 1 | US-08-232-018-5 |
| 4 | 19 | 100.0 | 41 | 1 | US-08-504-047-5 |
| 5 | 19 | 100.0 | 41 | 2 | US-09-087-855-5 |
| 6 | 19 | 100.0 | 48 | 2 | US-08-347-563A-23 |
| 7 | 19 | 100.0 | 48 | 3 | US-08-485-942A-23 |
| 8 | 19 | 100.0 | 48 | 3 | US-08-488-214A-23 |
| 9 | 19 | 100.0 | 48 | 3 | US-08-488-208A-23 |
| 10 | 19 | 100.0 | 48 | 4 | US-08-483-211A-23 |
| 11 | 19 | 100.0 | 48 | 4 | US-08-488-223A-23 |
| 12 | 19 | 100.0 | 48 | 4 | US-08-438-431A-23 |
| 13 | 19 | 100.0 | 53 | 4 | US-09-149-476-448 |
| 14 | 19 | 100.0 | 59 | 4 | US-09-088-651-5 |
| 15 | 19 | 100.0 | 72 | 3 | US-09-075-087-4 |
| 16 | 19 | 100.0 | 81 | 4 | US-08-924-629C-49 |
| 17 | 19 | 100.0 | 81 | 4 | US-08-924-629C-52 |
| 18 | 19 | 100.0 | 95 | 1 | US-08-436-265A-110 |
| 19 | 19 | 100.0 | 95 | 2 | US-08-971-217-110 |
| 20 | 19 | 100.0 | 95 | 4 | US-09-350-600-110 |
| 21 | 19 | 100.0 | 98 | 4 | US-09-134-001C-4312 |
| 22 | 19 | 100.0 | 99 | 4 | US-09-370-838-95 |
| 23 | 19 | 100.0 | 107 | 3 | US-08-540-406-2 |
| 24 | 19 | 100.0 | 107 | 3 | US-08-636-052-2 |
| 25 | 19 | 100.0 | 107 | 4 | US-08-954-668-2 |
| 26 | 19 | 100.0 | 107 | 4 | US-08-918-658-2 |
| 27 | 19 | 100.0 | 107 | 5 | PCT-US95-13233-2 |

| | | | | | | |
|----|----|-------|-----|---|-------------------|-------------------|
| 28 | 19 | 100.0 | 136 | 4 | US-09-325-932A-76 | Sequence 76, Appl |
| 29 | 19 | 100.0 | 160 | 1 | US-08-014-153D-8 | Sequence 8, Appl |
| 30 | 19 | 100.0 | 166 | 2 | US-08-347-563A-5 | Sequence 5, Appl |
| 31 | 19 | 100.0 | 166 | 2 | US-08-347-563A-6 | Sequence 6, Appl |
| 32 | 19 | 100.0 | 166 | 3 | US-08-292-345B-5 | Sequence 5, Appl |
| 33 | 19 | 100.0 | 166 | 3 | US-08-292-345B-6 | Sequence 6, Appl |
| 34 | 19 | 100.0 | 166 | 3 | US-08-485-942A-5 | Sequence 5, Appl |
| 35 | 19 | 100.0 | 166 | 3 | US-08-485-942A-6 | Sequence 6, Appl |
| 36 | 19 | 100.0 | 166 | 3 | US-08-488-214A-5 | Sequence 5, Appl |
| 37 | 19 | 100.0 | 166 | 3 | US-08-488-214A-6 | Sequence 6, Appl |
| 38 | 19 | 100.0 | 166 | 3 | US-08-488-208A-5 | Sequence 5, Appl |
| 39 | 19 | 100.0 | 166 | 3 | US-08-488-208A-6 | Sequence 6, Appl |
| 40 | 19 | 100.0 | 166 | 4 | US-08-692-922-2 | Sequence 2, Appl |
| 41 | 19 | 100.0 | 166 | 4 | US-08-483-211A-5 | Sequence 5, Appl |
| 42 | 19 | 100.0 | 166 | 4 | US-08-483-211A-6 | Sequence 6, Appl |
| 43 | 19 | 100.0 | 166 | 4 | US-08-488-223A-5 | Sequence 5, Appl |
| 44 | 19 | 100.0 | 166 | 4 | US-08-488-223A-6 | Sequence 6, Appl |
| 45 | 19 | 100.0 | 166 | 4 | US-08-438-431A-5 | Sequence 5, Appl |

ALIGNMENTS

RESULT 1
US-08-449-207-4
; Sequence 4, Application US/08449207
; Patent No. 5714313
; GENERAL INFORMATION:
; APPLICANT: Garfinkel, David J.
; APPLICANT: Nissley, Dwight V.
; APPLICANT: Curcio, Joan M.
; APPLICANT: Strathern, Jeffrey N.
; TITLE OF INVENTION: SIMPLE METHOD FOR DETECTING INHIBITORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, 127 Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: U.S.A.
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,207
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Selby, Elizabeth
; REGISTRATION NUMBER: 38,298
; REFERENCE/DOCKET NUMBER: 14014,0144
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-449-207-4

Query Match 100.0%; Score 19; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVPI 4
||||

Db 9 AVPI 12

RESULT 2

US-08-850-910A-20

Sequence 20, Application US/08850910A
Patent No. 5948761

GENERAL INFORMATION:

APPLICANT: SEILHAMER, J. J.

APPLICANT: LEWICKI, J.

APPLICANT: SCARBOROUGH, R. M.

TITLE OF INVENTION: RECOMBINANT TECHNIQUES FOR

TITLE OF INVENTION: PRODUCTION OF BRAIN NUTRIENTIC PEPTIDE

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESSES:

ADDRESSEE: MORRISON & FOERSTER, LLP

STREET: 2000 Pennsylvania Avenue, NW, Suite 5500

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1888

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/850,910A

FILING DATE: 05-MAY-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/477,226

FILING DATE: 08-FEB-1990

APPLICATION NUMBER: 07/299,880

FILING DATE: 19-JAN-1989

APPLICATION NUMBER: 07/206,470

FILING DATE: 14-JUN-1988

APPLICATION NUMBER: 07/200,383

FILING DATE: 31-MAY-1988

ATTORNEY/AGENT INFORMATION:

NAME: Mursahide, Kate H

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 219002025212

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-887-1500

TELEFAX: 202-822-0168

TELEX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

US-08-850-910A-20

OY 1 AVPI 4
1111
Db 2 AVPI 5

Query Match 100.0%; Score 19; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3

US-08-232-018-5

Sequence 5, Application US/08232018
Patent No. 5633158

GENERAL INFORMATION:

APPLICANT: Anlezark, Gillian M.

APPLICANT: Melton, Roger

APPLICANT: Sherwood, Roger
APPLICANT: Connors, Thomas
APPLICANT: Friedlos, Frank
APPLICANT: Jarman, Michael
APPLICANT: Knox, Richard
APPLICANT: Mauger, Anthony
TITLE OF INVENTION: Bacterial Nitroreductase for the
TITLE OF INVENTION: Reduction of CB 1954 and Analogues thereof to a Cytotoxic
TITLE OF INVENTION: Form
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Spencer, Frank & Schneider
STREET: Suite 300 East, 1100 New York Avenue, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20005-3955
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,018
FILING DATE: 27-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Scheller, John W.
REGISTRATION NUMBER: 26,031
REFERENCE/DOCKET NUMBER: KEMPJ 0017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 414-4000
TELEFAX: (202) 414-4040
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
US-08-232-018-5

Query Match 100.0%; Score 19; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
1111
Db 22 AVPI 25

RESULT 4

US-08-504-047-5

Sequence 5, Application US/08504047
Patent No. 5780585

GENERAL INFORMATION:

APPLICANT: Anlezark, Gillian M.

APPLICANT: Melton, Roger

APPLICANT: Sherwood, Roger

APPLICANT: Connors, Thomas

APPLICANT: Friedlos, Frank

APPLICANT: Jarman, Michael

APPLICANT: Knox, Richard

APPLICANT: Mauger, Anthony

TITLE OF INVENTION: Bacterial Nitroreductase for the

TITLE OF INVENTION: Reduction of CB 1954 and Analogues thereof to a Cytotoxic

TITLE OF INVENTION: Form

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Spencer, Frank & Schneider
STREET: Suite 300 East, 1100 New York Avenue, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20005-3935
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/504,047
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,018
FILING DATE: 27-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Schmeidler, John W.
REGISTRATION NUMBER: 26,031
REFERENCE/DOCKET NUMBER: KEMPJ 0017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 414-4000
TELEFAX: (202) 414-4040
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
US-08-504-047-5

Query Match 100.0%; Score 19; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
DB 22 AVPI 25

RESULT 5
US-09-087-855-5
Sequence 5, Application US/09087855
Patent No. 5977065
GENERAL INFORMATION:
APPLICANT: Anlezark, Gillian M.
APPLICANT: Melton, Roger
APPLICANT: Sherwood, Roger
APPLICANT: Connors, Thomas
APPLICANT: Friedlos, Frank
APPLICANT: Jarman, Michael
APPLICANT: Mauger, Anthony
TITLE OF INVENTION: Bacterial Nitroreductase for the
TITLE OF INVENTION: Reduction of CB 1954 and Analogues Thereof to a Cytotoxic
TITLE OF INVENTION: Form
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spencer, Frank & Schneider
STREET: Suite 300 East, 1100 New York Avenue, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20005-3935
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/087,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/232,018
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Schmeidler, John W.
REGISTRATION NUMBER: 26,031
REFERENCE/DOCKET NUMBER: KEMPJ 0017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 414-4000
TELEFAX: (202) 414-4040
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
US-09-087-855-5

Query Match 100.0%; Score 19; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
DB 22 AVPI 25

RESULT 6
US-08-347-563A-23
Sequence 23, Application US/08347563A
Patent No. 5935810
GENERAL INFORMATION:
APPLICANT: THE ROCKEFELLER UNIVERSITY
TITLE OF INVENTION: MODULATORS OF BODY WEIGHT, CORRESPONDING NUCLEIC
TITLE OF INVENTION: ACIDS AND PROTEINS, AND DIAGNOSTIC AND THERAPEUTIC USES THE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/347,563A
FILING DATE: NO. 5935810eember 30, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/292,345
FILING DATE: August 17, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-087 CIP
TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: N-terminal portion of the human ob protein
US-08-347-563A-23

Query Match
Best Local Similarity 100.0%; Score 19; DB 2; Length 48;
Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
||||
DB 21 AVPI 24

RESULT 7
US-08-485-942A-23
Sequence 23; Application US/08485942A
Patent No. 6048837
GENERAL INFORMATION:
APPLICANT: JEFFREY M. FRIEDMAN, YIYING ZHANG, RICARDO PROENCA,
APPLICANT: MARGHERITA MAFFEI, JEFFREY HALAAS, KETAN GAJIMWALA, AND STEPHEN K. BURLE
TITLE OF INVENTION: OB POLYPEPTIDE AS MODULATORS OF BODY WEIGHT (AS AMENDED)
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,942A
FILING DATE: JUNE 7, 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/438,431
FILING DATE: May 10, 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/347,563
FILING DATE: No. 6048837ember 30, 1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/292,345
FILING DATE: August 17, 1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-087 CIP 2F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

DESCRIPTION: N-terminal portion of the human ob protein
DESCRIPTION: encoded by first exon
US-08-485-942A-23

Query Match
Best Local Similarity 100.0%; Score 19; DB 3; Length 48;
Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
||||
DB 21 AVPI 24

RESULT 8
US-08-488-214A-23
Sequence 23; Application US/08488214A
Patent No. 6124439
GENERAL INFORMATION:
APPLICANT: JEFFREY M. FRIEDMAN, YIYING ZHANG, RICARDO PROENCA,
APPLICANT: MARGHERITA MAFFEI, JEFFREY HALAAS, KETAN GAJIMWALA, AND STEPHEN K. BU
TITLE OF INVENTION: OB POLYPEPTIDE ANTIBODIES AND METHOD OF MAKING
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,214A
FILING DATE: JUNE 7, 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/438,431
FILING DATE: May 10, 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/347,563
FILING DATE: No. 6124439ember 30, 1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/292,345
FILING DATE: August 17, 1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-087 CIP 2D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: N-terminal portion of the human ob protein
DESCRIPTION: encoded by first exon
US-08-488-214A-23

Query Match
Best Local Similarity 100.0%; Score 19; DB 3; Length 48;
Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
Db 21 AVPI 24

RESULT 9
US-08-488-208A-23
; Sequence 23, Application US/08488208A
; Patent No. 6124448
; GENERAL INFORMATION:
; APPLICANT: THE ROCKEFELLER UNIVERSITY
; TITLE OF INVENTION: MODULATORS OF BODY WEIGHT, CORRESPONDING
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS, AND DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,208A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/485,943
; FILING DATE: June 7, 1995
; APPLICATION NUMBER: 08/438,431
; FILING DATE: May 10, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/347,563
; FILING DATE: NO. 6124448ember 30, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/292,345
; FILING DATE: August 17, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-087 CIP2I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: N-terminal portion of the human ob protein
US-08-488-208A-23

Query Match 100.0%; Score 19; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVPI 4
Db 21 AVPI 24
RESULT 10
US-08-483-211A-23

; Sequence 23, Application US/08483211A
; Patent No. 6309853
; GENERAL INFORMATION:
; APPLICANT: THE ROCKEFELLER UNIVERSITY
; TITLE OF INVENTION: MODULATORS OF BODY WEIGHT, CORRESPONDING
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS, AND DIAGNOSTIC AND THERAPEUTIC
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,211A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/485,943
; FILING DATE: June 7, 1995
; APPLICATION NUMBER: 08/438,431
; FILING DATE: May 10, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/347,563
; FILING DATE: NO. 6309853ember 30, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/292,345
; FILING DATE: August 17, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-087 CIP2I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: N-terminal portion of the human ob protein
US-08-483-211A-23

Query Match 100.0%; Score 19; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AVPI 4
Db 21 AVPI 24
RESULT 11
US-08-488-223A-23
; Sequence 23, Application US/08488223A
; Patent No. 6350730
; GENERAL INFORMATION:
; APPLICANT: THE ROCKEFELLER UNIVERSITY
; TITLE OF INVENTION: MODULATORS OF BODY WEIGHT, CORRESPONDING NUCLEIC
; ACIDS AND PROTEINS, AND DIAGNOSTIC AND THERAPEUTIC USES
; NUMBER OF SEQUENCES: 98

```

CORRESPONDENCE ADDRESS:
ADDRESS: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,223A
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/485,943
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/347,563
FILING DATE: No. 6350730eember 30, 1994
APPLICATION NUMBER: 08/292,345
FILING DATE: August 17, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-087 CIP2I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: N-terminal portion of the human ob protein
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-08-488-223A-23

Query Match 100.0%; Score 19; DB 4; Length 48;
Match Local Similarity 100.0%; Pred. NO. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
| | |
Db 21 AVPI 24

RESULT 12
US-08-438-431A-23
Sequence 23, Application US/08438431A
Patent No. 6429290
GENERAL INFORMATION:
APPLICANT: JEFFREY M. FRIDMAN, YIYING ZHANG, RICARDO PROENCA, MARGHERITA MAFFEI
TITLE OF INVENTION: MODULATORS OF BODY WEIGHT, CORRESPONDING NUCLEIC ACIDS AND PR
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,431A
FILING DATE: May 10, 1995

```

```

1 CLASSIFICATION: 514
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: 08/347,563
4 FILING DATE: No. 6429290ember 30, 1994
5 CLASSIFICATION: 514
6 PRIOR APPLICATION DATA:
7 APPLICATION NUMBER: 08/292,345
8 FILING DATE: August 17, 1994
9 CLASSIFICATION: 514
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Jackson Esq., David A.
12 REGISTRATION NUMBER: 26,742
13 REFERENCE/DOCKET NUMBER: 600-1-087 CIP1
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: 201 487-5800
16 TELEFAX: 201 343-1684
17 TELEX: 133521
18 INFORMATION FOR SEQ ID NO: 23:
19 SEQUENCE CHARACTERISTICS:
20 LENGTH: 48 amino acids
21 TYPE: amino acid
22 TOPOLOGY: linear
23 MOLECULE TYPE: protein
24 DESCRIPTION: N-terminal portion of the human ob protein
25
26 US-08-438-431A-23
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
98
```

[illegible]

EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 100.0%; Score 19; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
1111
DB 36 AVPI 39

RESULT 14
US-09-088-651-5
Sequence 5, Application US/09088651
Patent No. 6165771

GENERAL INFORMATION:

APPLICANT: BURGESS, NICOLA A.
APPLICANT: CLINKENBEARD, HELEN E.
APPLICANT: SOUTHAN, CHRISTOPHER D.
TITLE OF INVENTION: NOVEL COMPOUNDS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA

ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/088,651
FILING DATE: JUNE 1, 1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB9712088.5
FILING DATE: 10-JUNE-1997

APPLICATION NUMBER: EP 97308295.1
FILING DATE: 17-OCT-1997

APPLICATION NUMBER: GB 9803650.2
FILING DATE: 20-FEB-1998

ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F.

REGISTRATION NUMBER: 23,031

REFERENCE/DOCKET NUMBER: GH30358

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-407-0700

TELEFAX: 610-407-0701

TELEX: 846169

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 59 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-088-651-5

Query Match 100.0%; Score 19; DB 4; Length 59;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
1111
DB 22 AVPI 25

RESULT 15
US-09-075-087-4
Sequence 4, Application US/09075087
Patent No. 6027932

GENERAL INFORMATION:

APPLICANT: Goldberg, Alfred L.

APPLICANT: Kandror, Olga

TITLE OF INVENTION: Methods of Improving Viability of Cells at Low Temperature

FILE REFERENCE: 10498/02966

CURRENT APPLICATION NUMBER: US/09/075,087

CURRENT FILING DATE: 1998-05-08

EARLIER APPLICATION NUMBER: 60/046,064

EARLIER FILING DATE: 1997-05-09

NUMBER OF SEQ ID NOS: 4

SOFTWARE: Patentlin Ver. 2.0

SEQ ID NO 4

LENGTH: 72

TYPE: PRT

ORGANISM: Escherichia coli

US-09-075-087-4

Query Match 100.0%; Score 19; DB 3; Length 72;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
1111
DB 14 AVPI 17

Search completed: February 16, 2003, 09:11:17
Job time : 39 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:09:28 ; Search time 28 Seconds
(without alignments)
3.650 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19
Sequence: 1 AVPI 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCY_NEW_PUB pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCYUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 19 | 100.0 | 4 | 9 US-10-068-569-21 | Sequence 21, Appl |
| 2 | 19 | 100.0 | 4 | 9 US-09-965-967-1 | Sequence 1, Appl |
| 3 | 19 | 100.0 | 4 | 10 US-09-939-293-13 | Sequence 13, Appl |
| 4 | 19 | 100.0 | 5 | 9 US-10-068-569-20 | Sequence 20, Appl |
| 5 | 19 | 100.0 | 5 | 9 US-10-068-569-24 | Sequence 24, Appl |
| 6 | 19 | 100.0 | 5 | 10 US-09-939-293-5 | Sequence 5, Appl |
| 7 | 19 | 100.0 | 7 | 9 US-10-068-569-12 | Sequence 12, Appl |
| 8 | 19 | 100.0 | 7 | 9 US-09-965-967-8 | Sequence 8, Appl |
| 9 | 19 | 100.0 | 7 | 10 US-09-939-293-6 | Sequence 6, Appl |
| 10 | 19 | 100.0 | 9 | 10 US-09-939-293-12 | Sequence 12, Appl |
| 11 | 19 | 100.0 | 10 | 9 US-09-965-967-18 | Sequence 18, Appl |
| 12 | 19 | 100.0 | 13 | 9 US-09-965-967-25 | Sequence 25, Appl |
| 13 | 19 | 100.0 | 15 | 9 US-10-068-569-8 | Sequence 8, Appl |
| 14 | 19 | 100.0 | 29 | 10 US-09-935-291A-42 | Sequence 42, Appl |
| 15 | 19 | 100.0 | 30 | 10 US-09-939-293-7 | Sequence 7, Appl |
| 16 | 19 | 100.0 | 35 | 10 US-09-939-293-11 | Sequence 11, Appl |
| 17 | 19 | 100.0 | 39 | 10 US-09-939-293-8 | Sequence 8, Appl |
| 18 | 19 | 100.0 | 40 | 10 US-09-939-293-2 | Sequence 2, Appl |
| 19 | 19 | 100.0 | 48 | 10 US-09-736-084-23 | Sequence 23, Appl |

| | | | | | |
|----|----|-------|-----|------------------------|--------------------|
| 20 | 19 | 100.0 | 54 | 10 US-09-864-761-41788 | Sequence 41788, A |
| 21 | 19 | 100.0 | 57 | 10 US-09-764-869-686 | Sequence 686, App |
| 22 | 19 | 100.0 | 62 | 9 US-09-796-692-1550 | Sequence 1550, App |
| 23 | 19 | 100.0 | 63 | 10 US-09-822-540A-1 | Sequence 1, Appl |
| 24 | 19 | 100.0 | 81 | 9 US-10-103-511-10 | Sequence 10, Appl |
| 25 | 19 | 100.0 | 81 | 10 US-09-864-761-46542 | Sequence 46542, A |
| 26 | 19 | 100.0 | 81 | 10 US-09-805-204-10 | Sequence 10, Appl |
| 27 | 19 | 100.0 | 84 | 10 US-09-798-116-9 | Sequence 9, Appl |
| 28 | 19 | 100.0 | 91 | 10 US-09-864-761-39729 | Sequence 39729, A |
| 29 | 19 | 100.0 | 99 | 9 US-09-854-133-95 | Sequence 95, Appl |
| 30 | 19 | 100.0 | 99 | 10 US-09-738-973-95 | Sequence 95, Appl |
| 31 | 19 | 100.0 | 102 | 10 US-09-925-297-633 | Sequence 633, App |
| 32 | 19 | 100.0 | 102 | 10 US-09-924-358-43 | Sequence 43, Appl |
| 33 | 19 | 100.0 | 114 | 10 US-09-263-959-320 | Sequence 320, App |
| 34 | 19 | 100.0 | 119 | 10 US-09-750-373-23 | Sequence 23, Appl |
| 35 | 19 | 100.0 | 149 | 9 US-10-125-452-25 | Sequence 25, Appl |
| 36 | 19 | 100.0 | 149 | 9 US-09-955-504-25 | Sequence 25, Appl |
| 37 | 19 | 100.0 | 166 | 10 US-09-736-084-5 | Sequence 5, Appl |
| 38 | 19 | 100.0 | 166 | 10 US-09-736-084-6 | Sequence 6, Appl |
| 39 | 19 | 100.0 | 166 | 10 US-09-932-888-2 | Sequence 2, Appl |
| 40 | 19 | 100.0 | 167 | 9 US-09-804-409A-11 | Sequence 11, Appl |
| 41 | 19 | 100.0 | 167 | 10 US-09-736-084-2 | Sequence 2, Appl |
| 42 | 19 | 100.0 | 167 | 10 US-09-736-084-4 | Sequence 4, Appl |
| 43 | 19 | 100.0 | 167 | 10 US-09-928-522-7 | Sequence 7, Appl |
| 44 | 19 | 100.0 | 167 | 10 US-09-928-522-8 | Sequence 8, Appl |
| 45 | 19 | 100.0 | 167 | 10 US-09-789-306-1 | Sequence 1, Appl |

ALIGNMENTS

RESULT 1
US-10-068-569-21
; Sequence 21, Application US/10068569
; Patent No. US20020160975A1
; GENERAL INFORMATION:
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandez-Alnemati, Teresa
; APPLICANT: Alnemati, Emdad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; FILE REFERENCE: 480140.475
; CURRENT APPLICATION NUMBER: US/10/068,569
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-068-569-21

Query Match 100.0%; Score 19; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 12e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 2
US-09-965-967-1
; Sequence 1, Application US/09965967
; Patent No. US2002017557A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Yigong
; TITLE OF INVENTION: Compositions and Methods for Regulating Apoptosis
; FILE REFERENCE: P0-0031 (01-1739-1)
; CURRENT APPLICATION NUMBER: US/09/965,967
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,574
; PRIOR FILING DATE: 2000-09-29

PRIOR APPLICATION NUMBER: 60/256,830
PRIOR FILING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 4
TYPE: PRT
ORGANISM: Homo sapiens
US-09-965-967-1

Query Match 100.0%; Score 19; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
Db 1 AVPI 4

RESULT 3
US-09-939-293-13
Sequence 13, Application US/09939293
Patent No. US20020132786A1
GENERAL INFORMATION:
APPLICANT: Alnemri, Emad S.
TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
TITLE OF INVENTION: AND METHODS OF USING THE SAME
FILE REFERENCE: 480140.465
CURRENT APPLICATION NUMBER: US/09/939,293
CURRENT FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 4
TYPE: PRT
ORGANISM: Homo sapiens
US-09-939-293-13

Query Match 100.0%; Score 19; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
Db 1 AVPI 4

RESULT 4
US-10-068-569-20
Sequence 20, Application US/10068569
Patent No. US20020160975A1
GENERAL INFORMATION:
APPLICANT: Srinivasula, Srinivasa M.
APPLICANT: Fernandes-Alnemri, Teresa
TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
TITLE OF INVENTION: CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS
FILE REFERENCE: 480140.475
CURRENT APPLICATION NUMBER: US/10/068,569
CURRENT FILING DATE: 2002-02-06
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 20
LENGTH: 5
TYPE: PRT
ORGANISM: Homo sapiens
US-10-068-569-20

Query Match 100.0%; Score 19; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4

Db 1 AVPI 4

RESULT 5
US-10-068-569-24
Sequence 24, Application US/10068569
Patent No. US20020160975A1
GENERAL INFORMATION:
APPLICANT: Srinivasula, Srinivasa M.
APPLICANT: Fernandes-Alnemri, Teresa
TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
TITLE OF INVENTION: CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS
FILE REFERENCE: 480140.475
CURRENT APPLICATION NUMBER: US/10/068,569
CURRENT FILING DATE: 2002-02-06
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 5
TYPE: PRT
ORGANISM: Homo sapiens
US-10-068-569-24

Query Match 100.0%; Score 19; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
Db 1 AVPI 4

RESULT 6
US-09-939-293-5
Sequence 5, Application US/09939293
Patent No. US20020132786A1
GENERAL INFORMATION:
APPLICANT: Alnemri, Emad S.
TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
TITLE OF INVENTION: AND METHODS OF USING THE SAME
FILE REFERENCE: 480140.465
CURRENT APPLICATION NUMBER: US/09/939,293
CURRENT FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 5
TYPE: PRT
ORGANISM: Homo sapiens
US-09-939-293-5

Query Match 100.0%; Score 19; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
Db 1 AVPI 4

RESULT 7
US-10-068-569-12
Sequence 12, Application US/10068569
Patent No. US20020160975A1
GENERAL INFORMATION:
APPLICANT: Srinivasula, Srinivasa M.
APPLICANT: Fernandes-Alnemri, Teresa
TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
TITLE OF INVENTION: CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS
FILE REFERENCE: 480140.475

;; CURRENT APPLICATION NUMBER: US/10/068,569
;; CURRENT FILING DATE: 2002-02-06
;; NUMBER OF SEQ ID NOS: 28
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 12
;; LENGTH: 7
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-068-569-12

Query Match 100.0%; Score 19; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
Db 1 AVPI 4

RESULT 8
US-09-965-967-8
;; Sequence 8, Application US/09965967
;; Patent No. US20020177557A1
;; GENERAL INFORMATION:
;; APPLICANT: Shi, Yigong
;; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
;; FILE REFERENCE: PU-0031 (01-1739-1)
;; CURRENT APPLICATION NUMBER: US/09/965,967
;; CURRENT FILING DATE: 2001-09-28
;; PRIOR APPLICATION NUMBER: 60/236,574
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/256,830
;; PRIOR FILING DATE: 2000-12-20
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 8
;; LENGTH: 7
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-967-8

Query Match 100.0%; Score 19; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
Db 1 AVPI 4

RESULT 9
US-09-939-293-6
;; Sequence 6, Application US/09939293
;; Patent No. US20020132786A1
;; GENERAL INFORMATION:
;; APPLICANT: Alnemri, Emad S.
;; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
;; FILE REFERENCE: 480140,465
;; CURRENT APPLICATION NUMBER: US/09/939,293
;; CURRENT FILING DATE: 2001-08-24
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 6
;; LENGTH: 7
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-939-293-6

Query Match 100.0%; Score 19; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
Db 1 AVPI 4

RESULT 10
US-09-939-293-12
;; Sequence 12, Application US/09939293
;; Patent No. US20020132786A1
;; GENERAL INFORMATION:
;; APPLICANT: Alnemri, Emad S.
;; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
;; FILE REFERENCE: 480140,465
;; CURRENT APPLICATION NUMBER: US/09/939,293
;; CURRENT FILING DATE: 2001-08-24
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 12
;; LENGTH: 9
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-939-293-12

Query Match 100.0%; Score 19; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
Db 5 AVPI 8

RESULT 11
US-09-965-967-18
;; Sequence 18, Application US/09965967
;; Patent No. US20020177557A1
;; GENERAL INFORMATION:
;; APPLICANT: Shi, Yigong
;; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
;; FILE REFERENCE: PU-0031 (01-1739-1)
;; CURRENT APPLICATION NUMBER: US/09/965,967
;; CURRENT FILING DATE: 2001-09-28
;; PRIOR APPLICATION NUMBER: 60/236,574
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/256,830
;; PRIOR FILING DATE: 2000-12-20
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 18
;; LENGTH: 10
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-967-18

Query Match 100.0%; Score 19; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
Db 1 AVPI 4

RESULT 12
US-09-965-967-25
;; Sequence 25, Application US/09965967
;; Patent No. US20020177557A1
;; GENERAL INFORMATION:
;; APPLICANT: Shi, Yigong
;; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
;; FILE REFERENCE: PU-0031 (01-1739-1)
;; CURRENT APPLICATION NUMBER: US/09/965,967

; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,574
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/256,830
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-965-967-25

Query Match 100.0%; Score 19; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
Db 4 AVPI 7

RESULT 13
US-10-068-569-8
; Sequence 8, Application US/10068569
; Patent No. US20020160975A1
; GENERAL INFORMATION:
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; FILE REFERENCE: 480140.475
; CURRENT APPLICATION NUMBER: US/10/068,569
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-068-569-8

Query Match 100.0%; Score 19; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
Db 1 AVPI 4

RESULT 14
US-09-935-291A-42
; Sequence 42, Application US/09935291A
; Patent No. US20020107373A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Chun, Myoung
; TITLE OF INVENTION: 49937, 49931, AND 49933, NOVEL HUMAN TRANSPORTER
; FILE REFERENCE: MNI-184
; CURRENT APPLICATION NUMBER: US/09/935,291A
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/226,504
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: 60/250,932
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 29
; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae
US-09-935-291A-42

Query Match 100.0%; Score 19; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
Db 14 AVPI 17

RESULT 15
US-09-939-293-7
; Sequence 7, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-7

Query Match 100.0%; Score 19; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
Db 1 AVPI 4

Search completed: February 16, 2003, 09:19:05
Job time : 29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:06:48 ; Search time 338 Seconds
(without alignments)
7.630 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

```
1: /cgn2_6/ptodata/2/paa/PCRS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 19 | 100.0 | 4 | 1 | PCT-US02-03553-21 |
| 2 | 19 | 100.0 | 4 | 1 | PCT-US02-17342-3 |
| 3 | 19 | 100.0 | 4 | 1 | PCT-US02-17342-34 |
| 4 | 19 | 100.0 | 4 | 1 | PCT-US02-17342-66 |
| 5 | 19 | 100.0 | 4 | 1 | PCT-US02-17342-68 |
| 6 | 19 | 100.0 | 4 | 20 | US-09-645-075-5 |

| | | | | | | |
|----|----|-------|----|----|--------------------|-------------------|
| 7 | 19 | 100.0 | 4 | 20 | US-09-687-549-14 | Sequence 14, Appl |
| 8 | 19 | 100.0 | 4 | 23 | US-09-939-293-13 | Sequence 13, Appl |
| 9 | 19 | 100.0 | 4 | 23 | US-09-965-967-1 | Sequence 1, Appl |
| 10 | 19 | 100.0 | 4 | 24 | US-10-068-569-21 | Sequence 21, Appl |
| 11 | 19 | 100.0 | 5 | 1 | PCT-US02-03553-20 | Sequence 20, Appl |
| 12 | 19 | 100.0 | 5 | 1 | PCT-US02-03553-24 | Sequence 24, Appl |
| 13 | 19 | 100.0 | 5 | 1 | PCT-US02-22658-2 | Sequence 2, Appl |
| 14 | 19 | 100.0 | 5 | 20 | US-09-645-075-6 | Sequence 6, Appl |
| 15 | 19 | 100.0 | 5 | 20 | US-09-687-549-9 | Sequence 9, Appl |
| 16 | 19 | 100.0 | 5 | 23 | US-09-939-293-5 | Sequence 5, Appl |
| 17 | 19 | 100.0 | 5 | 24 | US-10-068-569-20 | Sequence 20, Appl |
| 18 | 19 | 100.0 | 5 | 24 | US-10-068-569-24 | Sequence 24, Appl |
| 19 | 19 | 100.0 | 5 | 25 | US-10-197-634-2 | Sequence 2, Appl |
| 20 | 19 | 100.0 | 6 | 20 | US-09-645-075-7 | Sequence 7, Appl |
| 21 | 19 | 100.0 | 6 | 20 | US-09-687-549-8 | Sequence 8, Appl |
| 22 | 19 | 100.0 | 7 | 1 | PCT-US02-03553-12 | Sequence 12, Appl |
| 23 | 19 | 100.0 | 7 | 20 | US-09-645-075-8 | Sequence 8, Appl |
| 24 | 19 | 100.0 | 7 | 20 | US-09-687-549-7 | Sequence 7, Appl |
| 25 | 19 | 100.0 | 7 | 23 | US-09-939-293-6 | Sequence 6, Appl |
| 26 | 19 | 100.0 | 7 | 23 | US-09-965-967-8 | Sequence 8, Appl |
| 27 | 19 | 100.0 | 7 | 24 | US-10-068-569-12 | Sequence 12, Appl |
| 28 | 19 | 100.0 | 8 | 18 | US-09-412-863-6512 | Sequence 12, Ap |
| 29 | 19 | 100.0 | 8 | 20 | US-09-687-549-6 | Sequence 6, Appl |
| 30 | 19 | 100.0 | 9 | 1 | PCT-US00-12063-70 | Sequence 70, Appl |
| 31 | 19 | 100.0 | 9 | 1 | PCT-US02-17342-36 | Sequence 36, Appl |
| 32 | 19 | 100.0 | 9 | 17 | US-09-357-333-70 | Sequence 70, Appl |
| 33 | 19 | 100.0 | 9 | 20 | US-09-687-549-3 | Sequence 3, Appl |
| 34 | 19 | 100.0 | 9 | 20 | US-09-687-549-4 | Sequence 4, Appl |
| 35 | 19 | 100.0 | 9 | 20 | US-09-687-549-5 | Sequence 5, Appl |
| 36 | 19 | 100.0 | 9 | 20 | US-09-687-549-10 | Sequence 10, Appl |
| 37 | 19 | 100.0 | 9 | 20 | US-09-687-549-19 | Sequence 19, Appl |
| 38 | 19 | 100.0 | 9 | 23 | US-09-939-293-12 | Sequence 12, Appl |
| 39 | 19 | 100.0 | 10 | 18 | US-09-412-863-6514 | Sequence 6514, Ap |
| 40 | 19 | 100.0 | 10 | 20 | US-09-687-549-22 | Sequence 22, Appl |
| 41 | 19 | 100.0 | 10 | 23 | US-09-965-967-18 | Sequence 18, Appl |
| 42 | 19 | 100.0 | 12 | 22 | US-09-831-757-53 | Sequence 53, Appl |
| 43 | 19 | 100.0 | 13 | 23 | US-09-965-967-25 | Sequence 25, Appl |
| 44 | 19 | 100.0 | 15 | 1 | PCT-US00-32227-47 | Sequence 47, Appl |
| 45 | 19 | 100.0 | 15 | 1 | PCT-US02-03553-8 | Sequence 8, Appl |

ALIGNMENTS

```
RESULT 1
PCT-US02-03553-21
; Sequence 21, Application PC/TUS0203553
; GENERAL INFORMATION:
; APPLICANT: Thomas Jefferson University
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Alnemri, Emdad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; FILE REFERENCE: CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS
; CURRENT APPLICATION NUMBER: PCT/US02/03553
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-03553-21
```

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|---|------|---|
| QY | 1 | AVPI | 4 |
| DB | 1 | AVPI | 4 |

RESULT 2
PCT-US02-17342-3
; Sequence 3, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that B
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCT/US02/17342
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
PCT-US02-17342-3

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 1 AVPI 4

RESULT 3
PCT-US02-17342-34
; Sequence 34, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that B
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCT/US02/17342
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
; NAME/KEY: misc_feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: V is N-methylvaline
PCT-US02-17342-34

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 1 AVPI 4

RESULT 4
PCT-US02-17342-66
; Sequence 66, Application PC/TUS0217342
; GENERAL INFORMATION:

; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCT/US02/17342
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: I is N-Methylisoleucine
PCT-US02-17342-66

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 1 AVPI 4

RESULT 5
PCT-US02-17342-68
; Sequence 68, Application PC/TUS0217342
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Princeton University
; TITLE OF INVENTION: IAP Binding Peptides and Assays for Identifying Compounds that
; FILE REFERENCE: PU-0217
; CURRENT APPLICATION NUMBER: PCT/US02/17342
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/345,630
; PRIOR FILING DATE: 2002-01-03
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic tetrapeptide sequence
; NAME/KEY: misc_feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: V is N-methylvaline
PCT-US02-17342-68

Query Match 100.0%; Score 19; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 1 AVPI 4

RESULT 6
US-09-645-075-5
; Sequence 5, Application US/09645075
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong

```
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Apoptotic Compounds
; FILE REFERENCE: UTS0732
; CURRENT APPLICATION NUMBER: US/09/645,075
; CURRENT FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-645-075-5

Query Match          100.0%; Score 19; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 7
US-09-687-549-14
; Sequence 14, Application US/09687549
; GENERAL INFORMATION:
; APPLICANT: Fesik, Stephen W.
; APPLICANT: Betz, Stephen F.
; APPLICANT: Meadows, Robert P.
; APPLICANT: Liu, Zhihong
; APPLICANT: Olejniczak, Edward T.
; APPLICANT: Sun, Chaohong
; TITLE OF INVENTION: PEPTIDES FROM SMAC (DIABLO) AND METHODS
; FILE REFERENCE: 6742.US.01
; CURRENT APPLICATION NUMBER: US/09/687,549
; CURRENT FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide derived from smac
US-09-687-549-14

Query Match          100.0%; Score 19; DB 20; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 8
US-09-939-293-13
; Sequence 13, Application US/09939293
; GENERAL INFORMATION:
; APPLICANT: Alnemrl, Emed S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 4
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-293-13

Query Match          100.0%; Score 19; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 9
US-09-965-967-1
; Sequence 1, Application US/09965967
; GENERAL INFORMATION:
; APPLICANT: Shi, Yigong
; TITLE OF INVENTION: Compositions And Methods For Regulating Apoptosis
; FILE REFERENCE: PU-0031 (01-1739-1)
; CURRENT APPLICATION NUMBER: US/09/965,967
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,574
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/256,830
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln version 3.1
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-967-1

Query Match          100.0%; Score 19; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 10
US-10-068-569-21
; Sequence 21, Application US/10068569
; GENERAL INFORMATION:
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemrl, Teresa
; APPLICANT: Alnemrl, Emed S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; FILE REFERENCE: CASPASE-9 AND SMAC/DIABLO FOR MEDIATING APOPTOSIS
; CURRENT APPLICATION NUMBER: US/10/068,569
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-068-569-21

Query Match          100.0%; Score 19; DB 24; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
   ||||
Db 1 AVPI 4

RESULT 11
```

PCT-US02-03553-20
; Sequence 20, Application PC/TUS0203553
; GENERAL INFORMATION:
; APPLICANT: Thomas Jefferson University
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Alnemri, Emdad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; FILE REFERENCE: 480140.475PC
; CURRENT APPLICATION NUMBER: PCT/US02/03553
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-03553-20

Query Match 100.0%; Score 19; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 12
PCT-US02-03553-24
; Sequence 24, Application PC/TUS0203553
; GENERAL INFORMATION:
; APPLICANT: Thomas Jefferson University
; APPLICANT: Srinivasula, Srinivasa M.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Alnemri, Emdad S.
; TITLE OF INVENTION: A CONSERVED XIAP-INTERACTION MOTIF IN
; FILE REFERENCE: 480140.475PC
; CURRENT APPLICATION NUMBER: PCT/US02/03553
; CURRENT FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-03553-24

Query Match 100.0%; Score 19; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 13
PCT-US02-22658-2
; Sequence 2, Application PC/TUS0222658
; GENERAL INFORMATION:
; APPLICANT: Thomas Jefferson University
; APPLICANT: Alnemri, Emdad S.
; TITLE OF INVENTION: OXI AND DOMAINS THEREOF THAT DISRUPT
; FILE REFERENCE: 480140.479PC
; CURRENT APPLICATION NUMBER: PCT/US02/22658
; CURRENT FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2

; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-22658-2

Query Match 100.0%; Score 19; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 14
US-09-645-075-6
; Sequence 6, Application US/09645075
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunyang
; TITLE OF INVENTION: Apoptotic Compounds
; FILE REFERENCE: US0732
; CURRENT APPLICATION NUMBER: US/09/645,075
; CURRENT FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-645-075-6

Query Match 100.0%; Score 19; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||
Db 1 AVPI 4

RESULT 15
US-09-687-549-9
; Sequence 9, Application US/09687549
; GENERAL INFORMATION:
; APPLICANT: Fesik, Stephen W.
; APPLICANT: Belz, Stephen F.
; APPLICANT: Meadows, Robert P.
; APPLICANT: Liu, Zhihong
; APPLICANT: Olejniczak, Edward T.
; APPLICANT: Sun, Chaohong
; TITLE OF INVENTION: PEPTIDES FROM SMAC (DIABLO) AND METHODS
; FILE REFERENCE: 6742.US.01
; CURRENT APPLICATION NUMBER: US/09/687,549
; CURRENT FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide derived from smac
US-09-687-549-9

Query Match 100.0%; Score 19; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AVPI 4
| | |
Db 1 AVPI 4

Search completed: February 16, 2003, 09:17:08
Job time : 339 secs

THIS PAGE BLANK (USFO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:07:43 ; Search time 62 Seconds
(without alignments)
5.451 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 429898 seqs, 84487048 residues

Total number of hits satisfying chosen parameters: 429898

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents: AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_US02_37577-4
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 19 | 100.0 | 4 | PCT-US02-37577-4 | Sequence 4, Appl |
| 2 | 19 | 100.0 | 4 | US-10-302-811-4 | Sequence 4, Appl |
| 3 | 19 | 100.0 | 7 | PCT-US02-37577-5 | Sequence 5, Appl |
| 4 | 19 | 100.0 | 7 | US-10-302-811-5 | Sequence 5, Appl |
| 5 | 19 | 100.0 | 23 | US-09-084-303B-290 | Sequence 290, App |
| 6 | 19 | 100.0 | 23 | US-10-339-740-290 | Sequence 290, App |
| 7 | 19 | 100.0 | 51 | PCT-US02-32727-2663 | Sequence 2663, App |
| 8 | 19 | 100.0 | 51 | PCT-US02-32727-10214 | Sequence 10214, A |
| 9 | 19 | 100.0 | 51 | US-10-057-498-2663 | Sequence 2663, App |
| 10 | 19 | 100.0 | 51 | US-10-057-498-10214 | Sequence 10214, A |
| 11 | 19 | 100.0 | 54 | PCT-US02-32727-21353 | Sequence 21353, A |
| 12 | 19 | 100.0 | 54 | US-10-057-498-21353 | Sequence 21353, A |
| 13 | 19 | 100.0 | 59 | US-08-457-855E-10 | Sequence 10, Appl |
| 14 | 19 | 100.0 | 65 | PCT-US02-32727-19164 | Sequence 19164, A |
| 15 | 19 | 100.0 | 65 | US-10-057-498-19164 | Sequence 19164, A |
| 16 | 19 | 100.0 | 66 | US-09-513-999C-5045 | Sequence 5045, A |
| 17 | 19 | 100.0 | 72 | US-09-084-303B-199 | Sequence 199, App |
| 18 | 19 | 100.0 | 72 | US-10-339-740-199 | Sequence 199, App |
| 19 | 19 | 100.0 | 74 | US-10-276-781-1759 | Sequence 1759, App |
| 20 | 19 | 100.0 | 79 | PCT-US02-32727-9694 | Sequence 9694, App |
| 21 | 19 | 100.0 | 79 | US-10-057-498-9694 | Sequence 9694, App |
| 22 | 19 | 100.0 | 81 | PCT-US02-32727-15598 | Sequence 15598, A |
| 23 | 19 | 100.0 | 81 | US-10-057-498-15598 | Sequence 15598, A |
| 24 | 19 | 100.0 | 82 | PCT-US02-35606-90 | Sequence 90, Appl |
| 25 | 19 | 100.0 | 82 | PCT-US02-35606-131 | Sequence 131, Appl |
| 26 | 19 | 100.0 | 83 | PCT-US02-32727-7583 | Sequence 7583, App |

| | | | | | | |
|----|----|-------|----|---|----------------------|--------------------|
| 27 | 19 | 100.0 | 83 | 6 | US-10-057-498-7583 | Sequence 7583, App |
| 28 | 19 | 100.0 | 88 | 1 | PCT-US02-32727-1508 | Sequence 1508, App |
| 29 | 19 | 100.0 | 88 | 1 | PCT-US02-32727-11017 | Sequence 11017, A |
| 30 | 19 | 100.0 | 88 | 6 | US-10-057-498-1508 | Sequence 1508, App |
| 31 | 19 | 100.0 | 88 | 6 | US-10-057-498-11017 | Sequence 11017, A |
| 32 | 19 | 100.0 | 89 | 5 | US-09-513-999C-4408 | Sequence 4408, App |
| 33 | 19 | 100.0 | 89 | 5 | US-09-134-000C-5898 | Sequence 5898, App |
| 34 | 19 | 100.0 | 89 | 5 | US-09-134-000C-5898 | Sequence 5898, App |
| 35 | 19 | 100.0 | 91 | 6 | US-10-218-140-2280 | Sequence 2280, App |
| 36 | 19 | 100.0 | 95 | 1 | PCT-US02-32727-25823 | Sequence 25823, A |
| 37 | 19 | 100.0 | 95 | 1 | PCT-US02-32727-27544 | Sequence 27544, A |
| 38 | 19 | 100.0 | 95 | 5 | US-09-724-676A-92555 | Sequence 92555, A |
| 39 | 19 | 100.0 | 95 | 5 | US-10-057-498-25823 | Sequence 25823, A |
| 40 | 19 | 100.0 | 95 | 6 | US-10-057-498-27544 | Sequence 27544, A |
| 41 | 19 | 100.0 | 95 | 6 | US-10-057-498-27544 | Sequence 27544, A |
| 42 | 19 | 100.0 | 96 | 1 | PCT-US02-32727-8327 | Sequence 8327, App |
| 43 | 19 | 100.0 | 96 | 6 | US-10-057-498-8327 | Sequence 8327, App |
| 44 | 19 | 100.0 | 98 | 6 | US-10-092-411A-4312 | Sequence 4312, App |
| 45 | 19 | 100.0 | 99 | 5 | US-09-724-676-90885 | Sequence 90885, A |

ALIGNMENTS

RESULT 1
PCT-US02-37577-4
Sequence 4, Application PC/TUS0237577
GENERAL INFORMATION:
APPLICANT: The Burnham Institute
APPLICANT: Reed, John C.
APPLICANT: Houghten, Richard A.
APPLICANT: Neftci, Adel
APPLICANT: Ostresh, John M.
APPLICANT: Pinilla, Clemencia
APPLICANT: Welsh, Kate
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Depression of IAP-Inhibited Caspase
FILE REFERENCE: FP-LJ 5449
CURRENT APPLICATION NUMBER: PCT/US02/37577
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/331,957
PRIOR FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 4
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
PCT-US02-37577-4
Query Match
Best Local Similarity 100.0%; Score 19; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
DB 1 AVPI 4
US-10-302-811-4
Sequence 4, Application US/10302811
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Houghten, Richard A.
APPLICANT: Neftci, Adel
APPLICANT: Ostresh, John M.
APPLICANT: Pinilla, Clemencia
APPLICANT: Welsh, Kate
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Depression of IAP-Inhibited Caspase

```
FILE REFERENCE: P-LJ 5504
CURRENT APPLICATION NUMBER: US/10/302.811
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/331,957
PRIOR FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 4
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
US-10-302-811-4
```

```
Query Match          100.0%; Score 19; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AVPI 4
    ||||
Db 1 AVPI 4
```

```
RESULT 3
PCT-US02-37577-5
```

```
Sequence 5, Application PC/TUS0237577
GENERAL INFORMATION:
APPLICANT: The Burnham Institute
APPLICANT: Reed, John C.
APPLICANT: Houghten, Richard A.
APPLICANT: Neftzi, Adel
APPLICANT: Ostresh, John M.
APPLICANT: Pinilla, Clemencia
APPLICANT: Welsh, Kate
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Derepression of IAP-inhibited Caspase
FILE REFERENCE: P-LJ 5449
CURRENT APPLICATION NUMBER: PCT/US02/37577
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/331,957
PRIOR FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
FEATURE:
NAME/KEY: AMIDATION
LOCATION: 7
OTHER INFORMATION: at the C-terminus
FEATURE:
NAME/KEY: MOD_RES
LOCATION: 1
OTHER INFORMATION: hydrogenerated at the N-terminus
PCT-US02-37577-5
```

```
Query Match          100.0%; Score 19; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AVPI 4
    ||||
Db 1 AVPI 4
```

```
RESULT 4
US-10-302-811-5
```

```
Sequence 5, Application US/10302811
GENERAL INFORMATION:
```

```
APPLICANT: Reed, John C.
APPLICANT: Houghten, Richard A.
APPLICANT: Neftzi, Adel
APPLICANT: Ostresh, John M.
APPLICANT: Pinilla, Clemencia
APPLICANT: Welsh, Kate
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Derepression of IAP-inhibited Caspase
FILE REFERENCE: P-LJ 5504
CURRENT APPLICATION NUMBER: US/10/302.811
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/331,957
PRIOR FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
FEATURE:
NAME/KEY: AMIDATION
LOCATION: 7
OTHER INFORMATION: at the C-terminus
FEATURE:
NAME/KEY: MOD_RES
LOCATION: 1
OTHER INFORMATION: hydrogenerated at the N-terminus
US-10-302-811-5
```

```
Query Match          100.0%; Score 19; DB 6; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AVPI 4
    ||||
Db 1 AVPI 4
```

```
RESULT 5
US-09-084-303B-290
```

```
Sequence 290, Application US/09084303B
GENERAL INFORMATION:
APPLICANT: Doberstein, Stephen
APPLICANT: Reddy, Bindu
APPLICANT: Platt, Darren
APPLICANT: Ferguson, Kimberly
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF C. ELEGANS INSULIN-LIKE GENES AN
TITLE OF INVENTION: THEROOF
FILE REFERENCE: 7326-069-999
CURRENT APPLICATION NUMBER: US/09/084,303B
CURRENT FILING DATE: 1998-05-26
NUMBER OF SEQ ID NOS: 302
SOFTWARE: PatentIn version 3.1
SEQ ID NO 290
LENGTH: 23
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-084-303B-290
```

```
Query Match          100.0%; Score 19; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 AVPI 4
    ||||
Db 9 AVPI 12
```

```
RESULT 6
US-10-339-740-290
```

```
Sequence 290, Application US/10339740
```



```
; GENERAL INFORMATION:
; APPLICANT: Doberstein, Stephen
; APPLICANT: Reddy, Bindu
; APPLICANT: Platt, Darren
; APPLICANT: Ferguson, Kimberly
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF C. ELEGANS INSULIN-LIKE GENES AND T
; FILE REFERENCE: 7326-069-999
; CURRENT APPLICATION NUMBER: US/10/339,740
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US/09/084,303A
; PRIOR FILING DATE: 1998-05-26
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 290
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-339-740-290

Query Match          100.0%; Score 19; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
    ||||
Db 9 AVPI 12

RESULT 7
PCT-US02-32727-2663
; Sequence 2663, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhalla, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shylan
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 2663
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-2663

Query Match          100.0%; Score 19; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
    ||||
Db 18 AVPI 21

RESULT 8
PCT-US02-32727-10214
; Sequence 10214, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
```

```
; APPLICANT: Persing, David
; APPLICANT: Bhalla, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shylan
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 10214
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-10214

Query Match          100.0%; Score 19; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
    ||||
Db 13 AVPI 16

RESULT 9
US-10-057-498-2663
; Sequence 2663, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 2663
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
US-10-057-498-2663

Query Match          100.0%; Score 19; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
    ||||
Db 18 AVPI 21

RESULT 10
US-10-057-498-10214
; Sequence 10214, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 10214
; LENGTH: 51
```

TYPE: PRT
ORGANISM: Propionl acnes
US-10-057-498-10214

Query Match 100.0%; Score 19; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
1111
DB 13 AVPI 16

RESULT 11
PCT-US02-32727-21353
Sequence 21353, Application PC/TUS0232727
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
APPLICANT: Bhatia, Ajay
APPLICANT: Maisonneuve, Jean Francois
APPLICANT: Zhang, Yanni
APPLICANT: Wang, Siqing
APPLICANT: Jen, Shyian
APPLICANT: Lodes, Michael
APPLICANT: Benson, Darin
APPLICANT: Jones, Robert
APPLICANT: Carter, Darlick
APPLICANT: Barth, Brenda
APPLICANT: Douglass, John
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
FILE REFERENCE: 210121.514C1
CURRENT APPLICATION NUMBER: PCT/US02/32727
CURRENT FILING DATE: 2002-10-11
NUMBER OF SEQ ID NOS: 30992
SEQ ID NO 21353
LENGTH: 54
TYPE: PRT
ORGANISM: Propionl acnes
PCT-US02-32727-21353

Query Match 100.0%; Score 19; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
1111
DB 39 AVPI 42

RESULT 12
US-10-057-498-21353
Sequence 21353, Application US/10057498
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
FILE REFERENCE: 210121.514
CURRENT APPLICATION NUMBER: US/10/057,498
CURRENT FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 29212
SEQ ID NO 21353
LENGTH: 54
TYPE: PRT
ORGANISM: Propionl acnes
US-10-057-498-21353

Query Match 100.0%; Score 19; DB 6; Length 54;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
1111
DB 39 AVPI 42

RESULT 13
US-08-457-855E-10
Sequence 10, Application US/08457855E
GENERAL INFORMATION:
APPLICANT: Giese, Klaus W.
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: METHOD OF PRODUCTION OF AN ON OB POLYPEPTIDE IN A MAMMAL (as a
FILE REFERENCE: 1089.003
CURRENT APPLICATION NUMBER: US/08/457,855E
CURRENT FILING DATE: 1995-06-01
PRIOR APPLICATION NUMBER: US 08/437,834
PRIOR FILING DATE: 1995-05-08
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 59
TYPE: PRT
ORGANISM: Mus sp.
US-08-457-855E-10

Query Match 100.0%; Score 19; DB 4; Length 59;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
1111
DB 21 AVPI 24

RESULT 14
PCT-US02-32727-19164
Sequence 19164, Application PC/TUS0232727
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
APPLICANT: Bhatia, Ajay
APPLICANT: Maisonneuve, Jean Francois
APPLICANT: Zhang, Yanni
APPLICANT: Wang, Siqing
APPLICANT: Jen, Shyian
APPLICANT: Lodes, Michael
APPLICANT: Benson, Darin
APPLICANT: Jones, Robert
APPLICANT: Carter, Darlick
APPLICANT: Barth, Brenda
APPLICANT: Douglass, John
TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
FILE REFERENCE: 210121.514C1
CURRENT APPLICATION NUMBER: PCT/US02/32727
CURRENT FILING DATE: 2002-10-11
NUMBER OF SEQ ID NOS: 30992
SEQ ID NO 19164
LENGTH: 65
TYPE: PRT
ORGANISM: Propionl acnes
PCT-US02-32727-19164

Query Match 100.0%; Score 19; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
1111
DB 14 AVPI 17

RESULT 15

US-10-057-498-19164
; Sequence 19164, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 19164
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Propionibacterium
US-10-057-498-19164

Query Match 100.0%; Score 19; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
Db 14 AVPI 17

Search completed: February 16, 2003, 09:18:23
Job time : 63 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 09:03:28 ; Search time 50 Seconds
(without alignments)
7.691 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19

Sequence: 1 AVPI 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | length | ID | Description |
|------------|-------|-------------|--------|--------|---------------------|
| 1 | 19 | 100.0 | 40 | G98008 | O-acetylhomoserine |
| 2 | 19 | 100.0 | 50 | P00547 | capsid protein VP2 |
| 3 | 19 | 100.0 | 52 | A81042 | hypothetical prote |
| 4 | 19 | 100.0 | 55 | E95916 | hypothetical prote |
| 5 | 19 | 100.0 | 62 | G83991 | hypothetical prote |
| 6 | 19 | 100.0 | 63 | G96531 | hypothetical prote |
| 7 | 19 | 100.0 | 64 | S25961 | hypothetical prote |
| 8 | 19 | 100.0 | 67 | T16814 | FMRFamide-like pep |
| 9 | 19 | 100.0 | 69 | S17721 | H+-transporting tw |
| 10 | 19 | 100.0 | 70 | I39786 | H+-transporting tw |
| 11 | 19 | 100.0 | 70 | G84119 | ATP synthase subun |
| 12 | 19 | 100.0 | 71 | JN0362 | H+-transporting tw |
| 13 | 19 | 100.0 | 71 | I39784 | H+-transporting tw |
| 14 | 19 | 100.0 | 76 | T18179 | hypothetical prote |
| 15 | 19 | 100.0 | 77 | JC5645 | lymphotoxin beta - |
| 16 | 19 | 100.0 | 80 | JQ1284 | hypothetical 9.3K |
| 17 | 19 | 100.0 | 81 | E53589 | hypothetical prote |
| 18 | 19 | 100.0 | 87 | T33285 | hypothetical prote |
| 19 | 19 | 100.0 | 90 | A83323 | hypothetical prote |
| 20 | 19 | 100.0 | 92 | E47754 | ycenA protein homol |
| 21 | 19 | 100.0 | 93 | JF0055 | baseplate protein |
| 22 | 19 | 100.0 | 94 | D71284 | hypothetical prote |
| 23 | 19 | 100.0 | 95 | S76543 | hypothetical prote |
| 24 | 19 | 100.0 | 95 | T03700 | SARF 2d protein, T |
| 25 | 19 | 100.0 | 97 | C69321 | conserved hypothet |
| 26 | 19 | 100.0 | 100 | F86542 | integration host f |
| 27 | 19 | 100.0 | 100 | B72080 | integration host f |
| 28 | 19 | 100.0 | 100 | H71536 | probable integrati |
| 29 | 19 | 100.0 | 100 | F81691 | integration host f |

| | | | | | | |
|----|----|-------|-----|---|--------|--------------------|
| 30 | 19 | 100.0 | 100 | 2 | B47293 | hypothetical prote |
| 31 | 19 | 100.0 | 102 | 2 | D97331 | hypothetical prote |
| 32 | 19 | 100.0 | 106 | 2 | F71681 | hypothetical prote |
| 33 | 19 | 100.0 | 108 | 2 | E83765 | hypothetical prote |
| 34 | 19 | 100.0 | 108 | 2 | S01151 | hypothetical prote |
| 35 | 19 | 100.0 | 110 | 2 | S26279 | T-cell receptor be |
| 36 | 19 | 100.0 | 110 | 2 | B72295 | conserved hypothet |
| 37 | 19 | 100.0 | 110 | 2 | C72472 | hypothetical prote |
| 38 | 19 | 100.0 | 112 | 2 | D95269 | hypothetical prote |
| 39 | 19 | 100.0 | 113 | 2 | S17396 | T-cell receptor be |
| 40 | 19 | 100.0 | 113 | 2 | Af2744 | hypothetical prote |
| 41 | 19 | 100.0 | 114 | 2 | E97525 | hypothetical prote |
| 42 | 19 | 100.0 | 116 | 2 | A86769 | dihydroneopterin a |
| 43 | 19 | 100.0 | 118 | 2 | AE1739 | PTS system, cellob |
| 44 | 19 | 100.0 | 118 | 2 | I64028 | hypothetical prote |
| 45 | 19 | 100.0 | 120 | 2 | A69489 | LSU ribosomal prot |

ALIGNMENTS

RESULT 1

G98008
O-acetylhomoserine (thiol)-lyase (EC 4.2.99.10) [imported] - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001

C:Accession: G98008
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Leikowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; PMID:21429245; PMID:11544234

A:Accession: G98008

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-40 <R>

A:Cross-References: GB:AE007317; PIDN:AAK99899.1; PID:g15458720; GSPDB:GN00174

C:Genetics:

A:Gene: mety-truncation

C:Keywords: carbon-oxygen lyase

Query Match 100.0%; Score 19; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
Db 26 AVPI 29

RESULT 2

P00547
capsid protein VP24 - human herpesvirus 1 (fragments)

C:Species: human herpesvirus 1
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999

C:Accession: P00547
R:Davidson, M.D.; Rixon, F.J.; Davison, A.J.

J. Gen. Virol. 73, 2709-2713, 1992

A:Title: Identification of genes encoding two capsid proteins (VP24 and VP26) of herp

A:Reference number: P00544; PMID:93019027; PMID:1328483

A:Accession: P00547

A:Molecule type: protein

A:Residues: 1-19; 20-33; 34-50 <DAV>

A:Experimental source: strain 17

C:Genetics:

A:Gene: US26

C:Superfamily: varicella-zoster virus gene 33 protein

C:Keywords: capsid protein

Query Match 100.0%; Score 19; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||||

Db 8 AVPI 11

RESULT 3

hypotheetical protein NMB1795 [imported] - Neisseria meningitidis (strain MC58 serogroup A1042)

C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: A81042

R:Retcelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A. Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; et al.; Qiu, H.; Yamashiro, J.; Gill, J.; Scarlato, V.; Masiar, V.; Pizze, M.
Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; et al.
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: A81042

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-52 <RET>

A:Cross-references: GB:AE002530; GB:AE002098; NID:97227054; PIDN:AAF4213.1; PID:9722704

A:Experimental source: serogroup B, strain MC58

C:Genetics:
A:Gene: NMB1795

Query Match 100.0%; Score 19; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||||

Db 18 AVPI 21

RESULT 4

hypotheetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSymE E95916

C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: E95916

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernat Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo A:Reference number: A95842; MUID:21396508; PMID:11461431
A:Accession: E95916

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-55 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC48997.1; PID:915140482; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; et al.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation
C:Genetics:
A:Gene: Smb21024
A:Genome: Plasmid

Query Match 100.0%; Score 19; DB 2; Length 55;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||||

Db 48 AVPI 51

RESULT 5

hypotheetical protein BH2735 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G83991

R:Rakami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: G83991

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-62 <STO>

A:Cross-references: GB:AF001516; GB:BA000004; NID:910175192; PIDN:BA06454.1; GSPDB:G A:Experimental source: strain C-125

C:Genetics:
A:Gene: BH2735

Query Match 100.0%; Score 19; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||||

Db 41 AVPI 44

RESULT 6

hypotheetical protein F13F21.6 [imported] - Arabidopsis thaliana G96531

C:Species: Arabidopsis thaliana (mouse ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96531

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzla Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96531

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-63 <STO>

A:Cross-references: GB:AE005173; NID:95430751; PIDN:AMD43151.1; GSPDB:GN00141

C:Genetics:
A:Gene: F13F21.6
A:Map position: 1

Query Match 100.0%; Score 19; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||||

Db 54 AVPI 57

RESULT 7

hypotheetical protein 64 - liverwort (Marchantia polymorpha) mitochondrion S25961

C:Species: Marchantia polymorpha
C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 21-Jul-2000
C:Accession: S25961

R:Oda, K.; Yamato, K.; Ohta, E.; Nakamura, Y.; Takemura, M.; Nozato, N.; Akashi, K.;

J. Mol. Biol. 223, 1-7, 1992
A:Title: Gene organization deduced from the complete sequence of liverwort *Marchantia* pc
A:Reference number: S25941; MUID:92114051; PMID:1731062
A:Accession: S25961
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-64 <ODA>
A:Cross-references: EMBL:M68929; NID:g786182; PIDN:AAC09396.1; PID:g786183
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992
C:Genetics:
A:Genome: mitochondrion
C:Keywords: mitochondrion

Query Match 100.0%; Score 19; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
|||||
DB 39 AVPI 42

RESULT 8
16814
FKFamide-like peptide 10 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 31-Jan-2000
C:Accession: T16814; T42434
R.Wu, X.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of *C. elegans* cosmid T06C10.
A:Reference number: 218582
A:Accession: T16814
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-67 <WUX>
A:Cross-references: EMBL:U41013; NID:g1086629; PID:g1086631; PIDN:AAA82299.1; CESP:T06C1
R.Li, C.; Nelson, L.S.; Memmott, J.M.
submitted to the EMBL Data Library, January 1998
A:Reference number: 222162
A:Accession: T42434
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-67 <LIC>
A:Cross-references: EMBL:AF042396; PIDN:AAC08947.1
C:Genetics:
A:Gene: FLP-10; T06C10.4
A:Introns: 25/1

Query Match 100.0%; Score 19; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
|||||
DB 19 AVPI 22

RESULT 9
S17721
H+-transporting two-sector ATPase (EC 3.6.3.14) lipid-binding protein - *Bacillus firmus*
N:Alternate names: H+-transporting ATP synthase chain c
C:Species: *Bacillus firmus*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 14-Dec-2001
C:Accession: S17721
R.Ivey, D.M.; Krulwich, T.A.
Mol. Gen. Genet. 229, 292-300, 1991
A:Title: Organization and nucleotide sequence of the *atp* genes encoding the ATP synthase
A:Reference number: S17719; MUID:92017665; PMID:1833620
A:Accession: S17721
A:Molecule type: DNA
A:Residues: 1-69 <IVE>
A:Cross-references: EMBL:M60117; NID:g2988387; PIDN:AAC08039.1; PID:g142546

C:Genetics:
A:Gene: *atpE*
C:Superfamily: H+-transporting ATP synthase lipid-binding protein
C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; transmembrane p
F:1-26/Domain: transmembrane #status predicted <TM1>
F:54/Active site: Glu #status predicted

Query Match 100.0%; Score 19; DB 1; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
|||||
DB 55 AVPI 58

RESULT 10
139786
H+-transporting two-sector ATPase (EC 3.6.3.14) C chain - *Bacillus firmus*
C:Species: *Bacillus firmus*
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 03-Jun-2002
C:Accession: 139786
R.Ivey, D.M.; Krulwich, T.A.
Res. Microbiol. 143, 467-470, 1992
A:Title: Two unrelated alkaliphilic *Bacillus* species possess identical deviations fro
h the ATP.
A:Reference number: 139783; MUID:93079447; PMID:1448623
A:Accession: 139786
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-70 <RES>
A:Cross-references: GB:M64713; NID:g142568; PIDN:AAA22257.1; PID:g142570
C:Genetics:
A:Gene: *atpE*
C:Superfamily: H+-transporting ATP synthase lipid-binding protein
C:Keywords: hydrolase

Query Match 100.0%; Score 19; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
|||||
DB 55 AVPI 58

RESULT 11
G84119
ATP synthase subunit c *atpE* [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 17-May-2002
C:Accession: G84119
R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G84119
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-70 <STO>
A:Cross-references: GB:A001519; GB:BA000004; NID:g10176109; PIDN:BA07478.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: *atpE*
C:Superfamily: H+-transporting ATP synthase lipid-binding protein

Query Match 100.0%; Score 19; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
|||||
DB 55 AVPI 58

```
RESULT 12
JN0362
H+-transporting two-sector ATPase (EC 3.6.3.14) lipid-binding protein - Enterococcus fae
C:Species: Enterococcus faecalis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 14-Dec-2001
C:Accession: JN0362
R:Kocherginskaya, S.A.; Shakhparonov, M.I.; Aldanova, N.A.; Modyanov, N.N.; Ovchinnikov,
Bioorg. Khim. 8, 1569-1571, 1982
A:Title: Proton-translocating adenosinetriphosphatase from Streptococcus faecalis. Struc
A:Reference number: JN0362
A:Accession: JN0362
A:Molecule type: protein
A:Residues: 1-71 <KOC>
C:Superfamily: H+-transporting ATP synthase lipid-binding protein
C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; transmembrane prot
F:1-26/Domain: transmembrane #status predicted <TM1>
F:44-70/Domain: transmembrane #status predicted <TM2>
F:1/Modified site: N-formylmethionine #status experimental
F:54/Active site: Glu #status predicted

Query Match          100.0%; Score 19; DB 1; Length 71;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
    |||
Db 55 AVPI 58

RESULT 13
I39784
H+-transporting two-sector ATPase (EC 3.6.3.14) C chain - Bacillus alcalophilus
C:Species: Bacillus alcalophilus
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 03-Jun-2002
C:Accession: I39784
R:Riley, D.M.; Krulwich, T.A.
Res. Microbiol. 143, 467-470, 1992
A:Title: Two unrelated alkaliphilic Bacillus species possess identical deviations from s
h the ATP.
A:Reference number: I39783; MUID:93079447; PMID:1448623
A:Accession: I39784
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-71 <RES>
A:Cross-references: GB:M84712; NID:g142565; PIDN:AAA22255.1; PID:g142567
A:Genetics:
A:Gene: atpE
C:Superfamily: H+-transporting ATP synthase lipid-binding protein
C:Keywords: hydrolase

Query Match          100.0%; Score 19; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
    |||
Db 55 AVPI 58

RESULT 14
T18179
hypothetical protein a677L - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18179
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T18179
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
```

```
A:Residues: 1-76 <GRA>
A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC97051.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A:Note: a677L

Query Match          100.0%; Score 19; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
    |||
Db 10 AVPI 13

RESULT 15
JC5645
lymphotoxin beta - human
N:Alternate names: non-Hodgkin's lymphoma
C:Species: Homo sapiens (man)
C:Date: 28-Oct-1997 #sequence_revision 28-Oct-1997 #text_change 05-Nov-1999
C:Accession: JC5645
R:Warzocha, K.; Renard, N.; Charlot, C.; Bienvenu, J.; Collier, B.; Salles, G.
Biochem. Biophys. Res. Commun. 238, 273-276, 1997
A:Title: Identification of two lymphotoxin beta isoforms expressed in human lymphoid
A:Reference number: JC5645; MUID:97445965; PMID:9299492
A:Accession: JC5645
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-77 <MAR>
A:Cross-references: GB:U89922; NID:g1870795; PIDN:AAC51769.1; PID:g1870796
C:Comment: This protein is involved in the development and function of the immune sys

Query Match          100.0%; Score 19; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
    |||
Db 35 AVPI 38

Search completed: February 16, 2003, 09:10:26
Job time : 52 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2003, 07:53:37 ; Search time 38 seconds
(without alignments)
4.366 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19
Sequence: 1 AVPI 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------|---------------------|
| 1 | 19 | 100.0 | 59 | 1 YCXB_OENHO | Q9MTN0 oenothera h |
| 2 | 19 | 100.0 | 64 | 1 YM25_MARPO | P38467 marchantia h |
| 3 | 19 | 100.0 | 69 | 1 ATPPL_BACFI | P22483 bacillus fi |
| 4 | 19 | 100.0 | 71 | 1 ATPPL_BACAO | P25966 bacillus al |
| 5 | 19 | 100.0 | 71 | 1 ATPPL_ENTHR | P26682 enterococcu |
| 6 | 19 | 100.0 | 71 | 1 CERD_CERCA | O17513 ceratilis c |
| 7 | 19 | 100.0 | 81 | 1 YBM1_CARPI | P38583 carobacter |
| 8 | 19 | 100.0 | 94 | 1 Y753_TREPA | O83734 treponema p |
| 9 | 19 | 100.0 | 100 | 1 DBH_CHLPA | O928C7 chlamydia p |
| 10 | 19 | 100.0 | 100 | 1 DBH_CHLTR | O84269 chlamydia t |
| 11 | 19 | 100.0 | 106 | 1 Y266_RICPR | O928Q9 rickettsia |
| 12 | 19 | 100.0 | 108 | 1 DN04_RHORO | P15017 rhodospirill |
| 13 | 19 | 100.0 | 111 | 1 Y065_BACAN | O91M97 bacillus an |
| 14 | 19 | 100.0 | 114 | 1 GON2_TUPGB | O95336 tupia gils |
| 15 | 19 | 100.0 | 117 | 1 Y4NH_RHISN | P55580 rhizobium s |
| 16 | 19 | 100.0 | 118 | 1 YEL6_HAEIN | P44189 haemophilus |
| 17 | 19 | 100.0 | 120 | 1 RL24_ARCFU | O28365 archaeoglob |
| 18 | 19 | 100.0 | 125 | 1 CASK_LAMGU | Q28451 lama guanic |
| 19 | 19 | 100.0 | 133 | 1 S3AD_BACSU | P49781 bacillus su |
| 20 | 19 | 100.0 | 134 | 1 NRDI_YERPE | O82466 yerinia pe |
| 21 | 19 | 100.0 | 145 | 1 RRNA_YEAST | P38204 saccharomyc |
| 22 | 19 | 100.0 | 147 | 1 NUSG_METJA | O57818 methanococ |
| 23 | 19 | 100.0 | 163 | 1 Y7DB_RHIME | O33682 rhizobium m |
| 24 | 19 | 100.0 | 167 | 1 OB_BOVIN | P50595 bos taurus |
| 25 | 19 | 100.0 | 167 | 1 OB_CANFA | O02720 canis fami |
| 26 | 19 | 100.0 | 167 | 1 OB_FELCA | O9NAC1 felis silve |
| 27 | 19 | 100.0 | 167 | 1 OB_HUMAN | P41159 homo sapien |
| 28 | 19 | 100.0 | 167 | 1 OB_MACMU | Q28504 maceca mula |
| 29 | 19 | 100.0 | 167 | 1 OB_MOUSE | P41160 mus musculu |
| 30 | 19 | 100.0 | 167 | 1 OB_PIG | Q29406 sus scrofa |
| 31 | 19 | 100.0 | 167 | 1 OB_RAT | P50596 ratus norv |
| 32 | 19 | 100.0 | 167 | 1 OB_SMICR | O9XSW9 smitthopsis |
| 33 | 19 | 100.0 | 169 | 1 ATP2_MYCGA | P33250 mycoplasma |

| | | | | | |
|----|----|-------|-----|--------------|--------------------|
| 34 | 19 | 100.0 | 169 | 1 Y4TD_RHISN | P55658 rhizobium s |
| 35 | 19 | 100.0 | 171 | 1 VPF2_BBP2 | P22502 bacterioph |
| 36 | 19 | 100.0 | 171 | 1 Y4PD_RHISN | P55613 rhizobium s |
| 37 | 19 | 100.0 | 175 | 1 LPPP_MYCTU | P71882 mycobacteri |
| 38 | 19 | 100.0 | 176 | 1 Y1PC_YEAST | P53633 saccharomyc |
| 39 | 19 | 100.0 | 180 | 1 PYRE_METTH | O27888 methanobact |
| 40 | 19 | 100.0 | 182 | 1 CASK_CAMDR | P79139 camelus dro |
| 41 | 19 | 100.0 | 182 | 1 PYRE_PYRAB | P56814 pyrococcus |
| 42 | 19 | 100.0 | 182 | 1 PYRE_PYRBU | P58861 pyrococcus |
| 43 | 19 | 100.0 | 186 | 1 PYRE_PYRHO | O58855 pyrococcus |
| 44 | 19 | 100.0 | 189 | 1 OXYR_MYCXE | O87883 mycobacteri |
| 45 | 19 | 100.0 | 189 | 1 YHNN_CAEEL | Q94402 caenorhabd |

ALIGNMENTS

| | |
|------------|--|
| RESULT 1 | |
| YCXB_OENHO | |
| ID | YCXB_OENHO |
| AC | Q9MTN0; |
| DT | 15-JUN-2002 (Rel. 41, Created) |
| DT | 15-JUN-2002 (Rel. 41, Last sequence update) |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) |
| DE | Hypothetical 6.9 kDa protein in psbd-trn1 intergenic region (ORF59a). |
| OS | Oenothera hookeri (Hooker's evening primrose). |
| OG | Chloroplast. |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; |
| OC | eurosid II; Myrtales; Onagraceae; Oenothera. |
| OX | NCBI_TaxID=85636; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN=cv. Johansen; |
| RX | MEDLINE=20309318; PubMed=10852478; |
| RA | Hupfer H., Swiatek M., Hornung S., Herrmann R.G., Maier R.M., |
| RA | Chiu W.-L., Sears B.; |
| RT | *Complete nucleotide sequence of the Oenothera elata plastid |
| RT | chromosome, representing plastome I of the five distinguishable |
| RT | Eurothera plastomes.*; |
| RT | Mol. Gen. Genet. 263:581-585 (2000). |
| RL | ----- |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - |
| CC | the European Bioinformatics Institute. There are no restrictions on its |
| CC | use by non-profit institutions as long as its content is in no way |
| CC | modified and this statement is not removed. Usage by and for commercial |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ |
| CC | or send an email to license@sib-sib.ch). |
| CC | ----- |
| DR | EMBL: AJ271079; CAB67145.1; - |
| KW | Chloroplast; Hypothetical protein. |
| SQ | SEQUENCE 59 AA; 6877 MW; 5F74D8860109P63B CRC64; |
| Best Match | Score 19; DB 1; Length 59; |
| Query | 1 AVPI 4 |
| Db | 22 AVPI 25 |
| Matches | 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| RESULT 2 | |
| YCXB_OENHO | |
| ID | YM25_MARPO |
| AC | P38467; |
| DT | 01-OCT-1994 (Rel. 30, Created) |
| DT | 01-OCT-1994 (Rel. 30, Last sequence update) |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) |
| DE | Hypothetical 7.5 kDa protein in NAD5 5'region (ORF64). |
| GN | YMF25. |
| OS | Marchantia polymorpha (Liverwort). |

```

OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiaceae;
OC Marchantiaceae; Marchantia.
OX NCBI_TaxID=3197;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92114051; PubMed=1731062;
RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,
RA Akashi K., Kanegae T., Ogura Y., Kohchi T., Ohyama K.;
RT "Gene organization deduced from the complete sequence of liverwort
RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant
RT mitochondrial genome.";
RL J. Mol. Biol. 223:1-7(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M68929; AAC09396.1; -.
DR PIR: S25961; S25961.
KW Mitochondrion; Hypothetical protein.
SQ SEQUENCE 64 AA; 7473 MW; 14539C1A1B3D167D CRC64;

Query Match          100.0%; Score 19; DB 1; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
   |||
Db 39 AVPI 42

RESULT 3
ATPL_BACFI
ID ATPL_BACFI STANDARD: PRT; 69 AA.
AC P22483;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase C chain (EC 3.6.3.14) (lipid-binding protein).
GN ATPe.
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1399;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=OF4;
RX MEDLINE=92017665; PubMed=1833620;
RA Ivey D.M., Krulwich T.A.;
RT "Organization and nucleotide sequence of the atp genes encoding the
RT ATP synthase from alkaliphilic Bacillus firmus OF4.";
RL Mol. Gen. Genet. 229:293-300(1991).
CC -1- FUNCTION: THIS IS ONE OF THE THREE CHAINS OF THE NONENZYMATIC
CC COMPONENT (CF(0) SUBUNIT) OF THE ATPASE COMPLEX.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) = ADP + phosphate +
CC H(+)(out).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- MISCELLANEOUS: DICYLOHEXYLCARBODIIMIDE (DCCD) INHIBITS ATPASE.
CC -1- SIMILARITY: BELONGS TO THE ATPASE C CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF330160; AAC08039.1; -.
DR PIR: S17721; S17721.
DR HSSP: P00844; 1A91.
DR InterPro: IPR002379; ATPase_Csub.
DR InterPro: IPR000454; Eub_ATPase_Csub.
DR Pfam: PF00137; ATP-synt_C; 1.
DR TIGRfam: TIGR01260; ATP-synt_C; 1.
DR PROSITE: PS00605; ATPASE_C; 1.
KW Hydrogen ion transport; CF(0); Transmembrane; Lipid-binding.
FT BINDING 54 54 DICYLOHEXYLCARBODIIMIDE (BY SIMILARITY).
SQ SEQUENCE 69 AA; 6956 MW; 4B968F460EC3CF40 CRC64;

Query Match          100.0%; Score 19; DB 1; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
   |||
Db 55 AVPI 58

RESULT 4
ATPL_BACAO
ID ATPL_BACAO STANDARD: PRT; 71 AA.
AC P25966;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase C chain (EC 3.6.3.14) (lipid-binding protein).
GN ATPe.
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1445;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079447; PubMed=1448623;
RA Ivey D.M., Krulwich T.A.;
RT "Two unrelated alkaliphilic Bacillus species possess identical
RT deviations in sequence from those of other prokaryotes in regions of
RT F0 proposed to be involved in proton translocation through the ATP
RT synthase.";
RL Res. Microbiol. 143:467-470(1992).
CC -1- FUNCTION: THIS IS ONE OF THE THREE CHAINS OF THE NONENZYMATIC
CC COMPONENT (CF(0) SUBUNIT) OF THE ATPASE COMPLEX.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) = ADP + phosphate +
CC H(+)(out).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- MISCELLANEOUS: DICYLOHEXYLCARBODIIMIDE (DCCD) INHIBITS ATPASE.
CC -1- SIMILARITY: BELONGS TO THE ATPASE C CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M84712; AAA22255.1; -.
DR HSSP: P00844; 1A91.
DR InterPro: IPR002379; ATPase_Csub.
DR InterPro: IPR000454; Eub_ATPase_Csub.
DR Pfam: PF00137; ATP-synt_C; 1.
DR TIGRfam: TIGR01260; ATP-synt_C; 1.
DR PROSITE: PS00605; ATPASE_C; 1.
KW Hydrogen ion transport; CF(0); Transmembrane; Lipid-binding.
FT BINDING 54 54 DICYLOHEXYLCARBODIIMIDE (BY SIMILARITY).

```

SEQ SEQUENCE 71 AA: 7044 MW: 4ED0B301261B89FA CRC64;

Query Match 100.0%; Score 19; DB 1; Length 71;

Best Local Similarity 100.0%; Pred. No. 2e+02; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
||||
DB 55 AVPI 58

RESULT 5

ATPL_ENTHR STANDARD: PRT: 71 AA.

ID ATPL_ENTHR

AC P26682;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE ATP synthase C chain (EC 3.6.3.14) (Lipid-binding protein)

DE (Dicyclohexylcarbodimide-binding protein).

GN ATPe.

OS Enterococcus hirae.

OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.

OX NCBI_TaxID=1354;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 9790;

RX MEDLINE=93015650; PubMed=1328152;

RA Shibata C., Ehara T., Tomura K., Igarashi K., Kobayashi H.;

RT "Gene structure of Enterococcus hirae (Streptococcus faecalis) F1F0-

RT ATPase, which functions as a regulator of cytoplasmic pH.";

RL J. Bacteriol. 174:6117-6124(1992).

CC -1- FUNCTION: THIS IS ONE OF THE THREE CHAINS OF THE NONEZYMATIC

CC COMPONENT (CF(0) SUBUNIT) OF THE ATPASE COMPLEX.

CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) -> phosphate +

CC H(+)(out);

CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC

CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE

CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)

CC HAS THREE MAIN SUBUNITS: A, B AND C.

CC -1- MISCELLANEOUS: DICYCLOHEXYLCARBODIIMIDE (DCDI) INHIBITS ATPASE.

CC -1- SIMILARITY: BELONGS TO THE ATPASE C CHAIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL; M64265; AAA26861.1; -

DR EMBL; M90060; AAA26854.1; -

DR HSSP; P00844; 1A91.

DR InterPro: IPR002379; ATPase_Csub.

DR InterPro: IPR000454; Eub.ATPase_Csub.

DR Pfam: PF00137; ATP-synt_C; 1.

DR PRINTS: PR00124; ATPASEC.

DR TIGRPFAMS: TIGR01260; ATP_synt_c; 1.

DR PROSITE: PS00605; ATPASE_C; 1.

KW Hydrogen ion transport; CF(0); Transmembrane; Lipid-binding.

FT BINDING 54 54 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).

FT SEQUENCE 71 AA: 7294 MW: D1A15FF48C1651B CRC64;

Query Match 100.0%; Score 19; DB 1; Length 71;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
||||
DB 55 AVPI 58

RESULT 6

CERD CERCA

ID CERD.CERCA STANDARD: PRT: 71 AA.

AC 017513;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Ceratotoxin D precursor.

GN CTRXD.

OS Ceratitis capitata (Mediterranean fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Tephritidae; Tephritidae; Ceratitis.

OX NCBI_TaxID=7213;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Female accessory gland;

RX MEDLINE=98231103; PubMed=9569644;

RA Rosetto M., de Filippis T., Manetti A.G.O., Marchini D., Baldari C.T.,

RA Dalli R.;

RT "The genes encoding the antibacterial sex-specific peptides

RT ceratotoxins are clustered in the genome of the medfly *Ceratitis*

RT capitata.";

RL Insect Biochem. Mol. Biol. 27:1039-1046(1997).

CC -1- FUNCTION: FEMALE-SPECIFIC PEPTIDES WITH POTENT ACTIVITY AGAINST

CC GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA. THEY HAVE AS WELL

CC HEMOLYTIC ACTIVITY. THESE PROTEINS ARE STABLE EVEN AT 100 DEGREES

CC CELSIUS.

CC -1- SUBUNIT: HOMOPOLYMER OF FOUR TO SIX SUBUNITS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).

CC -1- SIMILARITY: STRUCTURALLY RELATED TO CECROPINS, DEFENSINS AND

CC APIADESINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL; Y15375; CAA75598.1; -

KW Insect immunity; Hemolysis; Antibiotic; Signal.

FT SIGNAL 1 23 POTENTIAL.

FT PROPEP 24 35 BY SIMILARITY.

FT PEPTIDE 36 71 CERATOTOXIN D.

FT SEQUENCE 71 AA: 7255 MW: 2B4B28E3D2848516 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 71;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
||||
DB 44 AVPI 47

RESULT 7
YBML_CARPI STANDARD: PRT: 81 AA.

ID YBML_CARPI

AC P38583;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 9.1 kDa protein in Bmi immunity protein 3'region.

OS Carnobacterium piscicola.

OC Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;

OC Carnobacterium.

OX NCBI_TaxID=2751;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LV17B;

RX MEDLINE=94216339; PubMed=8163526;

```

RA Quadri L.E.N., Sailer M., Roy K.L., Vederas J.C., Stiles M.E.;
RT "Chemical and genetic characterization of bacteriocins produced by
RL Carnobacterium discicola LV17B."
CC J. Biol. Chem. 269:12204-12211(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L29059; AAA72433.1;
KW Hypothetical protein.
SO SEQUENCE 81 AA; 9070 MW; 9FA00FB3441118FB CRC64;

Query Match 100.0%; Score 19; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
Db 37 AVPI 40

RESULT 8
ID Y753_TREPA STANDARD; PRT; 94 AA.
CD 083734;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0753.
GN TP0753.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nichols;
RC MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwyn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Soergers E., Hardson J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khelai H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001247; AAC65730.1;
DR TIGR: TP0753;
KW Hypothetical protein: Transmembrane; Complete proteome.
FT TRANSMEM 7 24 POTENTIAL.
FT TRANSMEM 39 61 POTENTIAL.
SO SEQUENCE 94 AA; 10330 MW; DEE0920E2A827877 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4

```

```

DB 39 AVPI 42

RESULT 9
ID DBH_CHLPN STANDARD; PRT; 100 AA.
AC 0928C7; 09J086;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable DNA-binding protein HU.
GN HUP OR CPN0416 OR CP0338.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CMT029;
RC MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AR39;
RC MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Baas S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwyn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-J138;
RC MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CML029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- FUNCTION: THIS PROTEIN BELONGS TO THE HISTONE LIKE FAMILY OF
CC PROKARYOTIC DNA-BINDING PROTEINS WHICH ARE CAPABLE OF WRAPPING
CC DNA TO STABILIZE IT, AND PREVENT ITS DENATURATION UNDER EXTREME
CC ENVIRONMENTAL CONDITIONS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL HISTONE-LIKE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001625; AAD18560.1;
DR EMBL: AE002196; AAF38192.1;
DR EMBL: AP002546; BAA98624.1;
DR HSSP: P08756; 1IHF.
DR TIGR: CP0338;
KW InterPro: IPR000119; Bac_DNABind.
KW Pfam: PF00216; Bac_DNA_binding.1.
DR PRODOM: PD000945; Bac_DNABind; 1.
DR SMART: SM00411; BHL; 1.
DR PROSITE: PS00045; HISTONE_LIKE; FALSR_NRG.
KW DNA-binding; DNA condensation; Complete proteome.
SO SEQUENCE 100 AA; 11474 MW; I700164BA575DDIC CRC64;

Query Match 100.0%; Score 19; DB 1; Length 100;

```

Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||||
DB 70 AVPI 73

RESULT 10

ID DBL_CHLTR STANDARD; PRT; 100 AA.
AC 084269;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN HUP OR CT267 OR TC0538.
OS Chlamydia trachomatis, and
OC Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxId=813, 83560;

RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=C.trachomatis; STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
RN [2]

RP SEQUENCE FROM N.A.
RC SPECIES=C.muridarum; STRAIN=Mopn / M199;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uitterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumonae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- FUNCTION: THIS PROTEIN BELONGS TO THE HISTONE LIKE FAMILY OF
CC PROKARYOTIC DNA-BINDING PROTEINS WHICH ARE CAPABLE OF WRAPPING
CC DNA TO STABILIZE IT, AND PREVENT ITS DENATURATION UNDER EXTREME
CC ENVIRONMENTAL CONDITIONS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL HISTONE-LIKE PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE001299; AAC67860.1; -;
DR EMBL: AE002322; AAF39378.1; -;
DR HSSP: P08756; 11HF.
DR TIGR: TC0538; -;
DR InterPro: IPR000119; Bac_DNABind.
DR Pfam: PF00216; Bac_DNA_binding.1.
DR ProDom: PDOM0945; Bac_DNABind; 1.
DR SMART: SM00411; BHL; 1.
DR PROSITE: PS00045; HISTONE_LIKE; FALSE_NEG.
DR DNA-binding; DNA condensation; Complete proteome.
KW SEQUENCE 100 AA; 11410 MW; 17446445F75C706 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||||
DB 70 AVPI 73

RESULT 11

ID Y266_RICPR STANDARD; PRT; 106 AA.
AC 092D09;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein RP266.
GN RP266.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxId=782;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,
RA Scharitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
RN [2]

CC -1- SIMILARITY: BELONGS TO THE UPF0091 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AJ235271; CA14728.1; -;
DR InterPro: IPR005133; Phag_Mnhg_YufB.
DR Pfam: PF03334; Phag_Mnhg_YufB; 1.
DR TIGRfams: TIGR01300; CPA3_mnhg_phag; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 106 AA; 11775 MW; FDDDB12887F165BF CRC64;

Query Match 100.0%; Score 19; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
|||||
DB 45 AVPI 48

RESULT 12

ID DNU4_RHQRU STANDARD; PRT; 108 AA.
AC P15017;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical transcriptional regulator in ATPase CF0 region (UNF4).
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC Rhodospirillum.
OX NCBI_TaxId=1085;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89025651; PubMed=2902844;
RA Falk G., Walker J.E.;
RT "DNA sequence of a gene cluster coding for subunits of the F0
membrane sector of ATP synthase in Rhodospirillum rubrum. Support for

RT Modular evolution of the F1 and F0 sectors."
RL Biochem. J. 254:109-122(1988).
CC -I- SIMILARITY: BELONGS TO THE PB5X(XRE) FAMILY OF TRANSCRIPTIONAL
CC REGULATORS. Y4AM/Y4MC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M37308; AAA26459.1; -
DR EMBL: X12757; CAA31250.1; -
DR PIR: S01151; S01151.
DR InterPro: IPR001387; HTH_3.
DR Pfam: PF01381; HTH_3; 1.
DR SMART: SM00530; HTH_XRE; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding.
FT DNA_BIND 31
FT H-T-H motif (PROBABLE).
SQ SEQUENCE 108 AA; 12489 MW; B0152E40FA673E07 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
DB 68 AVPI 71

RESULT 13
Y065_BACAN STANDARD; PRT; 111 AA.
AC O9RMW9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein pX02-65.
GN pX02-65.
OS Bacillus anthracis.
OC Plasmid pX02.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1392;
RN [1]
RP SEQUENCE FROM N.A.
RA Okinaka R.T., Cloud K., Hampton O., Hill K.K., Keim P., Lamke G.,
RA Kumano S., Mauter D., Martinez Y., Svensson R., Tatum L.R.,
RA Brown A.E., Jackson P.J.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF188935; AAF13670.1; -
KW Hypothetical protein; Plasmid; Transmembrane.
FT TRANSMEM 7
FT TRANSMEM 53 73
FT POTENTIAL.
SQ SEQUENCE 111 AA; 12601 MW; E6D778702CF4EB48 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
DB 15 AVPI 18

RESULT 14
G0N2_TUPGB STANDARD; PRT; 114 AA.
AC O95336;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin II precursor [Contains: Gonadoliberin II (LHRH II)
DE (luteinizing hormone releasing hormone II) (Gonadotropin releasing
DE hormone II) (GNRH II) (Luliberin II); GNRH-associated peptide II].
GN GNRH2.
OS Tupaiia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxId=9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal."
RT Gen. Comp. Endocrinol. 104:7-19(1996).
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: MIDBRAIN.
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U63327; AAB16838.1; -
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Signal.
FT SIGNAL 1 25
FT CHAIN 1 114
FT PEPTIDE 26 35
FT PEPTIDE 39 114
FT MOD_RES 26 26
FT MOD_RES 35 35
SQ SEQUENCE 114 AA; 12123 MW; 680E90E1C6869EC1 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
DB 109 AVPI 112

RESULT 15
Y4NH_RHISN STANDARD; PRT; 117 AA.
AC P55580;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 12.3 kDa protein Y4NH.

```

GN Y4NH.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: SOME, TO E. COLI AND S. AUREUS ETHIDIUM BROMIDE
CC RESISTANCE PROTEIN (EBR/OACC/EMRE/WVRC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000087; AAB91787.1; -.
DR InterPro: IPR000390; DUF7.
DR Pfam: PF00893; DUF7.1.
KW Hypothetical protein; Plasmid; Transmembrane.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 40 60 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
SQ SEQUENCE 117 AA; 12275 MW; 43AE4FC47A4E6847 CRC64;
Query Match 100.0%; Score 19; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AVPI 4
Db 3 AVPI 6

```

Search completed: February 16, 2003, 09:07:36
 Job time : 40 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: February 16, 2003, 08:57:08 ; Search time 94 Seconds
(without alignments)
8.768 Million cell updates/sec

Title: US-09-939-293-13

Perfect score: 19
Sequence: 1 AVPI 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database :

```

1:  sp.archaea:21:*
2:  sp.bacteria:1:*
3:  sp.fungi:1:*
4:  sp.human:1:*
5:  sp.invertebrate:*
6:  sp.mammal:1:*
7:  sp.mmc:1:*
8:  sp.organelle:*
9:  sp.plage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Score | Query | Match | Length | DB | ID | Description |
|--------|-------|-------|-------|--------|--------|--------|-------------|
| No. | | | | | | | |
| 1 | 19 | 100.0 | 27 | 13 | 09PR60 | 09PR60 | gadu morhu |
| 2 | 19 | 100.0 | 23 | 10 | 094HT2 | 094HT2 | oryza salti |
| 3 | 19 | 100.0 | 40 | 12 | 021K40 | 021K40 | hepatitis c |
| 4 | 19 | 100.0 | 41 | 5 | 08TR84 | 08TR84 | grosophila |
| 5 | 19 | 100.0 | 41 | 5 | 09Q582 | 09Q582 | human immun |
| 6 | 19 | 100.0 | 42 | 15 | 09Q582 | 09Q582 | human immun |
| 7 | 19 | 100.0 | 44 | 5 | 09NRP7 | 09NRP7 | ceratitis r |
| 8 | 19 | 100.0 | 47 | 16 | 0935M2 | 0935M2 | salmonella |
| 9 | 19 | 100.0 | 48 | 6 | 08SPQ4 | 08SPQ4 | sus scrofa |
| 10 | 19 | 100.0 | 48 | 15 | 091DS9 | 091DS9 | human immun |
| 11 | 19 | 100.0 | 49 | 15 | 09Q580 | 09Q580 | human immun |
| 12 | 19 | 100.0 | 49 | 15 | 09Q578 | 09Q578 | human immun |
| 13 | 19 | 100.0 | 50 | 12 | 065342 | 065342 | autographa |
| 14 | 19 | 100.0 | 52 | 16 | 091Y12 | 091Y12 | neisseria m |
| 15 | 19 | 100.0 | 54 | 12 | 09E3M1 | 09E3M1 | tomato yell |
| 16 | 19 | 100.0 | 55 | 16 | 092VY1 | 092VY1 | rhizobium m |
| | 19 | 100.0 | 59 | 8 | 09MTN0 | 09MTN0 | oenochera h |

| | | | | | | | |
|----|----|-------|----|----|--------|--------|--------------|
| 17 | 19 | 100.0 | 60 | 2 | 09X3F8 | 09X318 | prochlorococ |
| 18 | 19 | 100.0 | 60 | 10 | 09LJ63 | 09J163 | arabidopsis |
| 19 | 19 | 100.0 | 62 | 16 | 09K9B4 | 09K9D4 | bacillus ha |
| 20 | 19 | 100.0 | 63 | 10 | 09X1B7 | 09X1B7 | arabidopsis |
| 21 | 19 | 100.0 | 65 | 6 | 09TS93 | 09TS93 | bos taurus |
| 22 | 19 | 100.0 | 66 | 2 | 09K9X2 | 09K9X2 | mycoplasma |
| 23 | 19 | 100.0 | 67 | 5 | 022242 | 022242 | caenorhabdi |
| 24 | 19 | 100.0 | 68 | 12 | 09Z602 | 09Z602 | tomato yell |
| 25 | 19 | 100.0 | 68 | 12 | 09YKD4 | 09YKD4 | tomato yell |
| 26 | 19 | 100.0 | 68 | 12 | 08UZC0 | 08UZC0 | tomato bego |
| 27 | 19 | 100.0 | 68 | 12 | 08UZB7 | 08UZB7 | tomato bego |
| 28 | 19 | 100.0 | 68 | 12 | 08UZB7 | 08UZB7 | tomato bego |
| 29 | 19 | 100.0 | 70 | 2 | 059201 | 059201 | caenorhabdi |
| 30 | 19 | 100.0 | 70 | 16 | 09K6H0 | 09K6H0 | bacillus ha |
| 31 | 19 | 100.0 | 71 | 10 | 094EA1 | 094EA1 | oryza sativ |
| 32 | 19 | 100.0 | 72 | 15 | 090U59 | 090U59 | human |
| 33 | 19 | 100.0 | 72 | 15 | 090U53 | 090U53 | human |
| 34 | 19 | 100.0 | 75 | 15 | 090U60 | 090U60 | human |
| 35 | 19 | 100.0 | 75 | 15 | 090U58 | 090U58 | human |
| 36 | 19 | 100.0 | 75 | 15 | 090U57 | 090U57 | human |
| 37 | 19 | 100.0 | 75 | 15 | 090U56 | 090U56 | human |
| 38 | 19 | 100.0 | 75 | 15 | 090U55 | 090U55 | human |
| 39 | 19 | 100.0 | 75 | 15 | 090U54 | 090U54 | human |
| 40 | 19 | 100.0 | 75 | 15 | 090U52 | 090U52 | human |
| 41 | 19 | 100.0 | 75 | 15 | 090U51 | 090U51 | human |
| 42 | 19 | 100.0 | 75 | 15 | 090U50 | 090U50 | human |
| 43 | 19 | 100.0 | 75 | 15 | 090U49 | 090U49 | human |
| 44 | 19 | 100.0 | 75 | 15 | 090U48 | 090U48 | human |
| 45 | 19 | 100.0 | 75 | 15 | 090U47 | 090U47 | human |

ALIGNMENTS

RESULT 1

| | | | |
|--------|--------------|------|-------|
| ID | PRELIMINARY; | PRT; | 23 AA |
| Q9PRS0 | | | |

| | | |
|--|---|---|
| DP | 01-MAY-2000 | (TREMBLrel. 13, Created) |
| DT | 01-MAY-2000 | (TREMBLrel. 13, Last sequence update) |
| DT | 01-DEC-2001 | (TREMBLrel. 19, Last annotation update) |
| DE | Chymotrypsin isoenzyme CHT2 (Fragment). | |
| OS | Gadus morhua (Atlantic cod). | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; | |
| OC | Acanthomorphi; Paracanthopterygii; Gadiformes; Gadidae; Gadus. | |
| OX | NCBI_TaxID=8049; | |
| RP | [1] | |
| RN | SEQUENCE. | |
| RP | MEDLINE=96043258; PubMed=7584866; | |
| RX | Rae A.J., Flensburg R., Sletten K.; | |
| RT | "Chymotrypsin isoenzymes in Atlantic cod: differences in kinetics and | |
| RT | substrate specificity"; | |
| RL | Comp. Biochem. Physiol. 112B:393-398(1995). | |
| FT | NON_TER 1 | |
| FT | NON_TER 23 | |
| SO | SEQUENCE 23 AA; 2356 MW; A5EC2A7B73EB3F5E CRC64; | |
| Query Match 100.0%; Score 19; DB 13; Length 23; | | |
| Best Local Similarity 100.0%; Pred. No. 4.4e+02; | | |
| Matches | 4; Conservative | 0; Mismatches |
| | | 0; Indels |
| | | 0; Gaps |
| QY | 1 AVPI 4 | |
| | | |
| Db | 20 AVPI 23 | |
| RESULT 2 | | |
| Q94HU2 | | |
| Q94HU2 | PRELIMINARY; | PRT; 27 AA. |
| AC | Q94HU2; | |
| DT | 01-DEC-2001 | (TREMBLrel. 19, Created) |
| DT | 01-DEC-2001 | (TREMBLrel. 19, Last sequence update) |

```

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 3.2 Kda protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrharioideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsing Y.-I.C., Chow T.-Y., Wu H.-P., Chao Y.-T., Lee P.-F., Su C.-L.,
RA Liu S.-M., Chen C.-S., Shaw J.-F.;
RT "Oryza sativa PAC P0574H01 genomics sequence.";
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC079022; AK73152.1; -.
KW Hypothetical protein.
SQ SEQUENCE 27 AA; 3167 MW; A1C3DFBC4C316AE1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 10; Length 27;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
    |||
DB 18 AVPI 21

RESULT 3
O91K40 PRELIMINARY; PRT; 40 AA.
ID O91K40:
AC O91K40:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Genome polyprotein (Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)) (Fragment).
DE Hepatitis C virus.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H5;
RA Gimenez-Barcons M., Franco S., Suarez Y., Forns X., Ampurdanes S.,
RA Puig-Basagolt F., Sanchez-Fuero A., Barrera J.M., Llovet J.M.,
RA Bruix J., Sanchez-Tapias J.M., Rodes J., Saz J.C.;
RT "High amino acid variability within the NS5A gene of hepatitis C virus
RT (HCV) is associated with hepatocellular carcinoma in patients with
RT HCV-1b related cirrhosis.";
RL Hepatology 0:0-0(2001).
DR EMBL; AF379232; AA63449.1; -.
DR InterPro; IPR002531; HCV_NS1.
DR Pfam; PF01560; HCV_NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER
FT NON_TER
SQ SEQUENCE 40 AA; 3957 MW; 6A97AB578B0706C6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 12; Length 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
    |||
DB 1 AVPI 4

RESULT 4
O8T3R4 PRELIMINARY; PRT; 41 AA.
ID O8T3R4:
AC O8T3R4:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

```

```

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE AT08987P.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brockstein P., Hong L., Abhayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celnik S.;
RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY094637; AA10990.1; -.
SQ SEQUENCE 41 AA; 4474 MW; F06286F2F6D2619D CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 5; Length 41;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
    |||
DB 12 AVPI 15

RESULT 5
O9Q582 PRELIMINARY; PRT; 42 AA.
ID O9Q582:
AC O9Q582:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Envelope glycoprotein V2 region (Fragment).
DE ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang B., Saksena N.K.;
RT "HIV-1 Strains from a cohort of American subjects reveal the presence
RT of a V2 region extension unique to slow progressors and non-
RT progressors.";
RL AIDS 0:0-0(2000).
DR EMBL; AF203211; AA24360.1; -.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER
FT NON_TER
SQ SEQUENCE 42 AA; 4790 MW; DE78892C9F92A38B CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 15; Length 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
    |||
DB 25 AVPI 28

RESULT 6
O9NFP7 PRELIMINARY; PRT; 44 AA.
ID O9NFP7:
AC O9NFP7:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ceratotoxin 1 precursor (Fragment).
GN CRI.
OS Ceratitis rosa.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Tephritidae; Tephritidae; Ceratitis.
 OC NCBI_TaxID=56958;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rosetto M.;
 RT "Evolution of the ceratotoxin gene family in the medfly *Ceratitis capitata* and the Natal fruit fly *Ceratitis rosa*."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ272450; CAB75957.1; -.
 KW Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 17 POTENTIAL.
 FT CHAIN 30 >44 CERATOTOXIN 1.
 FT NON_TER 44 44
 SQ SEQUENCE 44 AA; 4676 MW; C81B7D0CDD4AB270 CRC64;
 Query Match 100.0%; Score 19; DB 5; Length 44;
 Best Local Similarity 100.0%; Pred. No. 8.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVPI 4
 DB 38 AVPI 41
 RESULT 7
 ID 0935M2 PRELIMINARY; PRT; 47 AA.
 AC 0935M2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein HCM1.141ac.
 GN HCM1.141ac.
 GN Salmonella typhi.
 OS Salmonella typhi.
 OC Plasmid PHCM1.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OC NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.F.G., Sebalhia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Conerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*
 RT enterica serovar Typhi CT18."
 RL Nature 413:848-852(2001).
 DR EMBL: AL513383; CAD09739.1; -.
 KW Hypothetical protein; Plasmid; Complete proteome.
 SQ SEQUENCE 47 AA; 5152 MW; 432B31C74096757 CRC64;
 Query Match 100.0%; Score 19; DB 16; Length 47;
 Best Local Similarity 100.0%; Pred. No. 8.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVPI 4
 DB 28 AVPI 31
 RESULT 8
 ID 08SP04 PRELIMINARY; PRT; 48 AA.
 AC 08SP04;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Leptin (Fragment).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Soares M.A.M., Euclydes R.F., Guimaraes S.E.F., Martins M.F.,
 RA Lopes P.S.;
 RT "Sus scrofa leptin gene (obese gene)."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY079082; AAL86020.1; -.
 FT NON_TER 48 48
 SQ SEQUENCE 48 AA; 5753 MW; 3053E5791FD6C5D CRC64;
 Query Match 100.0%; Score 19; DB 6; Length 48;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVPI 4
 DB 21 AVPI 24
 RESULT 9
 ID 09IDS9 PRELIMINARY; PRT; 48 AA.
 AC 09IDS9;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Envelope protein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OC NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IND.CAL684;
 RA Mandal D., Jana S., Panda S., Bhattacharya S., Ghosh T.C.,
 RA Bhattacharya S.K., Chakrabarti S.;
 RT "Distribution of HIV-1 subtypes in female sex workers in Calcutta,
 RT India."
 RL Indian J. Med. Res. 112:165-172(2000).
 DR EMBL: AJ278374; CAB93888.1; -.
 KW AIDS; Coat protein; Glycoprotein.
 FT NON_TER 1 1
 FT NON_TER 48 48
 SQ SEQUENCE 48 AA; 5383 MW; F0C0A0C7576D4E5 CRC64;
 Query Match 100.0%; Score 19; DB 15; Length 48;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVPI 4
 DB 4 AVPI 7
 RESULT 10
 ID 09Q580 PRELIMINARY; PRT; 49 AA.
 AC 09Q580;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Envelope glycoprotein V2 region (Fragment).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OS Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OC NCBI_TaxID=11676;

```

RN [1]
RP SEQUENCE FROM N.A.
RA Wang B., Saksena N.K.;
RT "HIV-1 strains from a cohort of American subjects reveal the presence
of a V2 region extension unique to slow progressors and non-
progressors."
RL AIDS 0:0-0(2000).
DR EMBL; AF203213; AAF24362.1; -.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1 1
FT NON_TER 49 49
SQ SEQUENCE 49 AA; 5572 MW; 0C08C08EF1AB29A3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 15; Length 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 25 AVPI 28

```

```

RESULT 11
Q90578
ID Q90578 PRELIMINARY; PRT; 49 AA.
AC Q90578;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DR 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Envelope glycoprotein V2 region (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang B., Saksena N.K.;
RT "HIV-1 strains from a cohort of American subjects reveal the presence
of a V2 region extension unique to slow progressors and non-
progressors."
RL AIDS 0:0-0(2000).
DR EMBL; AF203215; AAF24364.1; -.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1 1
FT NON_TER 49 49
SQ SEQUENCE 49 AA; 5527 MW; 0C0398EC21AB29A3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 15; Length 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 25 AVPI 28

```

```

RT "Regulation of expression of a baculovirus ecdysteroid UDPglucosyl-
transferase gene."
RT J. Virol. 64:1321-1328(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-L-1;
RX MEDLINE=93267802; PubMed=8497062;
RA Passarelli A.L., Miller L.K.;
RT "Identification and characterization of lef-1, a baculovirus gene
involved in late and very late gene expression."
RL J. Virol. 67:3481-3488(1993).
DR EMBL; L09723; AAA46708.1; -.
SQ SEQUENCE 50 AA; 5675 MW; 14606AA410E83C5D CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 12; Length 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 6 AVPI 9

```

```

RESULT 13
Q9JY12
ID Q9JY12 PRELIMINARY; PRT; 52 AA.
AC Q9JY12;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein NMB1795.
GN NMB1795.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MC58 / SEROGRUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
Hart D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Yamathavan J.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappunli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RT Science 287:1809-1815(2000).
DR EMBL; AE002529; AAF42133.1; -.
DR TIGR; NMB1795; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 52 AA; 6018 MW; 280F5A2090A82D7B CRC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 16; Length 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVPI 4
DB 18 AVPI 21

```

```

RESULT 14
Q9E3W1
ID Q9E3W1 PRELIMINARY; PRT; 54 AA.
AC Q9E3W1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Coat protein (Fragment).
GN VI.

```

OS Tomato yellow leaf curl virus (TYLCV).
 OC Viruses: ssDNA viruses; Geminiviridae; Begomovirus.
 OX NCBI_TaxID=10832;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ying Z., Davis M.J.;
 RT "Partial characterization and host range of tomato yellow leaf curl
 virus in south Florida."
 RL Proc. Annu. Meet. Fla. State Hort. Soc. 0:0-0(2000).
 DR EMBL; AF260331; AAC23626.1; -.
 FT NON_TER 54 54
 SQ SEQUENCE 54 AA; 6349 MW; A626B63B92E07376 CRC64;

Query Match 100.0%; Score 19; DB 12; Length 54;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
 ||||
 Db 32 AVPI 35

RESULT 15

O92VVI PRELIMINARY; PRT; 55 AA.
 ID O92VVI
 AC O92VVI:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein RB0597.
 GN RB0597 OR SMB21024.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSymb (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorhoevelter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
 RA Golding B., Puehler A.;
 RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
 fixing endosymbiont Sinorhizobium meliloti."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AL603644; CAC48997.1; -.
 KW Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 55 AA; 5986 MW; 250DEDB85B81943 CRC64;

Query Match 100.0%; Score 19; DB 16; Length 55;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVPI 4
 ||||
 Db 48 AVPI 51

Search completed: February 16, 2003, 09:09:23
 Job time : 96 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 03:15:41 ; Search time 3187 Seconds

(Without alignments)
767.065 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84
Sequence: 1 actcattcttcagtcacag.....gtttccagattgataaga 84

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 965098

Minimum DB seq length: 0

Maximum DB seq length: 84

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pla:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 23.6 | 28.1 | 65 | 6 | AX485057 |
| 2 | 20 | 23.6 | 71 | 14 | AF371357 |
| 3 | 19.8 | 23.6 | 80 | 6 | A73092 |
| 4 | 19.8 | 23.6 | 80 | 6 | A73184 |
| 5 | 19.8 | 23.6 | 80 | 6 | AR126972 |
| 6 | 19.8 | 23.6 | 82 | 3 | AY096255 |
| 7 | 19.6 | 23.6 | 65 | 6 | AX483305 |
| 8 | 19.4 | 23.1 | 83 | 6 | AX241055 |
| 9 | 19.2 | 22.9 | 60 | 5 | XELTRNIC1 |
| 10 | 18.8 | 22.4 | 81 | 3 | AF121763 |
| 11 | 18.8 | 22.4 | 81 | 3 | AF121764 |
| 12 | 18.6 | 22.1 | 61 | 9 | HUMBCL5D |
| 13 | 18.6 | 22.1 | 73 | 5 | AF454277 |
| 14 | 18.6 | 22.1 | 83 | 6 | A81641 |
| 15 | 18.4 | 21.9 | 61 | 6 | AX043859 |
| 16 | 18.4 | 21.9 | 61 | 6 | AX360855 |
| 17 | 18.4 | 21.9 | 61 | 6 | AX428223 |
| 18 | 18.4 | 21.9 | 71 | 14 | AF371356 |
| 19 | 18.4 | 21.9 | 71 | 14 | AF371358 |
| 20 | 18.4 | 21.9 | 71 | 14 | AF371360 |
| 21 | 18.4 | 21.9 | 81 | 3 | AF371363 |
| 22 | 18.4 | 21.9 | 81 | 3 | AF121765 |
| 23 | 18.4 | 21.9 | 81 | 3 | AF121766 |
| 24 | 18.4 | 21.9 | 81 | 3 | AF363236 |
| 25 | 18.2 | 21.7 | 64 | 3 | D28746 |
| 26 | 18.2 | 21.7 | 73 | 5 | AF454275 |
| 27 | 18 | 21.4 | 39 | 6 | AR116809 |
| 28 | 18 | 21.4 | 51 | 6 | AX074130 |
| 29 | 18 | 21.4 | 65 | 6 | AX483087 |
| 30 | 18 | 21.4 | 80 | 6 | AR204272 |
| 31 | 17.8 | 21.2 | 62 | 4 | AF164071 |
| 32 | 17.8 | 21.2 | 65 | 6 | AX484799 |
| 33 | 17.8 | 21.2 | 77 | 5 | AF051709 |
| 34 | 17.8 | 21.2 | 80 | 10 | MMDBDS9 |
| 35 | 17.8 | 21.2 | 82 | 5 | AF051712 |
| 36 | 17.6 | 21.0 | 75 | 6 | AR099225 |
| 37 | 17.6 | 21.0 | 80 | 10 | MM0403426 |
| 38 | 17.6 | 21.0 | 81 | 14 | AB015324 |
| 39 | 17.4 | 20.7 | 31 | 6 | AR201830 |
| 40 | 17.4 | 20.7 | 54 | 6 | AR196731 |
| 41 | 17.4 | 20.7 | 63 | 10 | MMVIMV13 |
| 42 | 17.2 | 20.5 | 65 | 6 | AX483821 |
| 43 | 17.2 | 20.5 | 65 | 6 | AX485361 |
| 44 | 17.2 | 20.5 | 67 | 9 | HSGPOCA2 |
| 45 | 17.2 | 20.5 | 82 | 14 | AF290014 |

ALIGNMENTS

| RESULT 1 | AX485057 | 65 bp | DNA | linear | PAT 16-AUG-2002 |
|------------|--|-------|-----|--------|-----------------|
| LOCUS | AX485057 | | | | |
| DEFINITION | Sequence 2357 from Patent WO02053728. | | | | |
| ACCESSION | AX485057 | | | | |
| VERSION | AX485057.1 | | | | |
| KEYWORDS | GI:22319341 | | | | |
| SOURCE | | | | | |
| ORGANISM | Candida albicans. | | | | |
| REFERENCE | | | | | |
| AUTHORS | Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L. | | | | |
| TITLE | Gene disruption methodologies for drug target discovery | | | | |
| JOURNAL | Patent: WO 02053728-A 2357 11-JUL-2002. | | | | |

Pred. No. is the number of results predicted by chance to have a

FEATURES Elitra Pharmaceuticals, Inc. (US)
Location/Qualifiers
source 1..65
/organism="Candida albicans"
/db_xref="taxon:5476"
BASE COUNT 21 a 3 c 16 g 25 t
ORIGIN
Query Match 28.1%; Score 23.6; DB 6; Length 65;
Best Local Similarity 64.8%; Pred. No. 4.8e+03;
Matches 35; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 14 GGTACAGACAGTGTGTGTCCTGCTGCTTGAAGACGCGTGT 67
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 GGTAAAGACATGTGTGTGTGTGATTCATCATAGTATTAATGAGAT 58
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 2 AF371357 71 bp RNA linear VRL 17-JUN-2001
LOCUS AF371357
DEFINITION Human metapneumovirus isolate 93-3 nucleoprotein (N) gene, partial
cDS
ACCESSION AF371357 GI:14485157
VERSION AF371357.1
KEYWORDS Human metapneumovirus.
SOURCE Human metapneumovirus.
ORGANISM Human metapneumovirus.
viruses; ssRNA negative-strand viruses; Mononegavirales;
Paramyxoviridae; Pneumovirinae; Metapneumovirus.
REFERENCE 1 (bases 1 to 71)
AUTHORS van den Hoogen,B.G., de Jong,J.C., Groen,J., Kuiken,T., de
Groen,R., Fouchier,R.A. and Osterhaus,A.D.
TITLE A newly discovered human pneumovirus isolated from young children
with respiratory tract disease
JOURNAL Nat. Med. 7 (6), 719-724 (2001)
MEDLINE 21279106
PUBMED 11385510
REFERENCE 2 (bases 1 to 71)
AUTHORS van den Hoogen,B.G., de Jong,J.C., Groen,J., Kuiken,T., de
Groen,R., Fouchier,R.A. and Osterhaus,A.D.
TITLE Direct Submission
JOURNAL Submitted (18-APR-2001) Dept of Virology, Erasmus University, Dr.
Molewaterplein 50, Rotterdam 3015 GE, The Netherlands
FEATURES
source 1..71
/organism="Human metapneumovirus"
/isolate="93-3"
/db_xref="taxon:162145"
gene <1..>71
/gene="N"
CDS <1..>71
/gene="N"
/codon_start=1
/product="nucleoprotein"
/protein_id="AAK62959.1"
/db_xref="GI:14485158"
/translation="RSFYDLFEQKYVHRSLEFYGKA"
BASE COUNT 24 a 10 c 13 g 24 t
ORIGIN
Query Match 23.8%; Score 20; DB 14; Length 71;
Best Local Similarity 65.9%; Pred. No. 6.7e+04;
Matches 29; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 34 GTTCCGTGTTGGCTACTTAAGACGCGTGTTCACAGATT 77
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 GATCCTTCTATGACTTATTGACAAAGTGTATCACAGAGT 45
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 3 A73092 80 bp DNA linear PAT 15-OCT-1999
LOCUS A73092
DEFINITION Sequence 68 from Patent WO9429451.

ACCESSION A73092
VERSION A73092.1 GI:6063964
KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 80)
AUTHORS Morgan,S.A. and Emtage,J.S.
TITLE HUMANISED ANTIBODIES
JOURNAL Patent: WO 9429451-A 68 22-DEC-1994;
CELLTECH LTD (GB); MORGAN SUSAN ADRIENNE (GB)
FEATURES
source 1..80
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 22 a 26 c 20 g 12 t
ORIGIN
Query Match 23.6%; Score 19.8; DB 6; Length 80;
Best Local Similarity 69.2%; Pred. No. 7.7e+04;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 28 TTGTGTCTCCTGTCGTGCTGCTTGAAGACGCGTGT 66
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48 TTGGGTCTTCCTGTCCTCTCTGTCGTGACACAGAGT 10
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 4 A73184 80 bp DNA linear PAT 15-OCT-1999
LOCUS A73184
DEFINITION Sequence 68 from Patent WO9429351.
ACCESSION A73184
VERSION A73184.1 GI:6064039
KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 80)
AUTHORS Morgan,S.A. and Emtage,J.S.
TITLE ANTIBODIES
JOURNAL Patent: WO 9429351-A 68 22-DEC-1994;
CELLTECH LTD (GB); MORGAN SUSAN ADRIENNE (GB)
FEATURES
source 1..80
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 22 a 26 c 20 g 12 t
ORIGIN
Query Match 23.6%; Score 19.8; DB 6; Length 80;
Best Local Similarity 69.2%; Pred. No. 7.7e+04;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 28 TTGTGTCTCCTGTCGTGCTGCTTGAAGACGCGTGT 66
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 48 TTGGGTCTTCCTGTCCTCTCTGTCGTGACACAGAGT 10
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 5 ARI26972 80 bp DNA linear PAT 16-MAY-2001
LOCUS ARI26972
DEFINITION Sequence 68 from patent US 6180377.
ACCESSION ARI26972
VERSION ARI26972.1 GI:14113565
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 80)
AUTHORS Morgan,S.Adrienne., Emtage,J.Spencer., Bodmer,M.William. and
Attnwal,D.Singh.
TITLE Humanized antibodies
JOURNAL Patent: US 6180377-A 68 30-JAN-2001;

| | | | | | |
|--|---|---------------------------------------|----|--|--|
| FEATURES | | | | | |
| SOURCE | Location/Qualifiers 1..80 /organism="unknown" | | | | |
| BASE COUNT | 22 a 26 c 20 g 12 t | | | | |
| ORIGIN | | | | | |
| Query Match 23.6% Score 19.8; DB 6; Length 80; Best Local Similarity 69.2%; Pred. No.7.7e+04; Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0; | | | | | |
| OY | 28 | TTGTGTCCTGTTGGCTACCTTTAGAACGCCTG | 66 | | |
| Dn | 48 | TTGGGTGTCCTGTCCTTCTCTGTCACACAGCACT | 10 | | |
| RESULT 6 | | | | | |
| LOCUS | AY096255 | 82 bp DNA linear INV 19-JUN-2002 | | | |
| DEFINITION | Chrysosora quinguecitrha homeobox protein neoscxl.Cc (neoscxl) | | | | |
| ACCESSION | AY096255 | | | | |
| VERSION | AY096255 | | | | |
| KEYWORDS | AY096255.1 GI:21206578 | | | | |
| SOURCE | Chrysosora quinguecitrha. | | | | |
| ORGANISM | Chrysosora quinguecitrha. Eukaryota; Metazoa; Cnidaria; Scyphozoa; Semeostomeae; Pelagidae; Chysosora. | | | | |
| REFERENCE | 1 (bases 1 to 82) Krong,G., Burton,P., Mazza,M., Ryan,J., Chaga,O. and Finnerty,J.R. Homeobox Evolution In Basal Animals: Reconstructing the Ancestral Emmetazoan Unpublished | | | | |
| AUTHORS | 2 (bases 1 to 82) Krong,G., Burton,P., Mazza,M. and Finnerty,J.R. | | | | |
| JOURNAL | Direct Submission | | | | |
| REFERENCE | Submitted (15-APR-2002) Department of Biology, Boston University, 5 Cunnington Street, Boston, MA 02215, USA | | | | |
| TITLE | location/Qualifiers | | | | |
| JOURNAL | 1..82 | | | | |
| FEATURES | /organism="Chrysosora quinguecitrha" /db_xref="taxon:6148" complement(<1..>82) /gene="neoscxl" /note="neoscxl.Cc" complement(<1..>82) /gene="neoscxl" /product="homeobox protein neoscxl.Cc" complement(<1..>82) /gene="neoscxl" /codon_start=1 /product="homeobox protein neoscxl.Cc" /protein_id="AA33448.1" /db_xref="GI:21206579" /translation="LYSRYLYTKERKDLSETLHTEROIKI" | | | | |
| CDS | BASE COUNT 16 a 13 c 15 g 38 t ORIGIN | | | | |
| Query Match 23.6% Score 19.8; DB 3; Length 82; Best Local Similarity 69.2%; Pred. No.7.6e+04; Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0; | | | | | |
| OY | 39 | TGTTGGCTACCTTTAGAACCGGTGTTCTCAGAATT | 77 | | |
| Dn | 7 | TGATTTGTCCTTCTGTAGATGCAGCGTTTCTGATAAT | 45 | | |
| RESULT 7 | | | | | |
| LOCUS | AX483305 | 65 bp DNA linear PAT 16-AUG-2002 | | | |
| DEFINITION | Sequence 605 from Patent WO02053728. | | | | |
| ACCESSION | AX483305 | | | | |
| VERSION | AX483305.1 GI:22317725 | | | | |
| KEYWORDS | | | | | |

| | |
|-----------------------|--|
| SOURCE | Candida albicans. |
| ORGANISM | Candida albicans. |
| REFERENCE | Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida. |
| AUTHORS | 1 Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L. |
| JOURNAL | Gene disruption methodologies for drug target discovery |
| DOI | Patent: WO 02053728-A 605 11-JUL-2002; |
| FEATURES | Elittra Pharmaceuticals, Inc. (US) |
| SOURCE | Location/Qualifiers |
| BASE COUNT | 1. .65 |
| ORIGIN | /organism="Candida albicans" |
| | /db_xref="taxon:3476" |
| Query Match | 23.3%; Score 19.6; DB 6; Length 65; |
| Best Local Similarity | 66.7%; Pred. No. 9.2e+04; |
| Matches | 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0; |
| Oy | 3 TTCATTCTTCAGGTACACAGACTGTTGGTGGTCTGTTGT 44 |
| Db | 47 TTCTCTCTCTTGTGTACACATTATTAATTATTCACACTGT 6 |
| RESULT 8 | |
| LOCUS | AX241055 83 bp DNA linear PAT 26-SEP-2001 |
| DEFINITION | Sequence 293 from Patent WO0160975. |
| ACCESSION | AX241055 |
| VERSION | AX241055.1 GI:15797930 |
| KEYWORDS | . |
| SOURCE | synthetic construct. |
| ORGANISM | artificial sequences. |
| REFERENCE | 1 (bases 1 to 83) |
| AUTHORS | Roemer,T., Jiang,B., Boone,C. and Bussey,H. |
| TITLE | Gene disruption methodologies for drug target discovery |
| JOURNAL | Patent: WO 0160975-A 293 23-AUG-2001; |
| FEATURES | Elittra Pharmaceuticals, Inc. (US) |
| SOURCE | Location/Qualifiers |
| BASE COUNT | 1. .83 |
| ORIGIN | /organism="synthetic construct" |
| | /db_xref="taxon:32630" |
| | /note="DNA primer" |
| Query Match | 23.1%; Score 19.4; DB 6; Length 83; |
| Best Local Similarity | 55.1%; Pred. No. 1e+05; |
| Matches | 38; Conservative 0; Mismatches 31; Indels 0; Gaps 0; |
| Oy | 8 TCTTCAGGTACACAGAGTGTGTGTTCTCTGTGTGGCTTAATTAGAAGCGGTGT 67 |
| Db | 4 TATTTCTTTAATGTAATGTGTGTGTGTGATCGTATTAATTGGAAGAATGAT 63 |
| Oy | 68 TCTCAGAT 76 |
| Db | 64 TGTGATTAT 72 |
| RESULT 9 | |
| LOCUS | XETRN1C1 60 bp DNA linear VRT 28-APR-1993 |
| DEFINITION | X.laavis transposon Txl, left junction B4. |
| ACCESSION | M13091 |
| VERSION | M13091.1 GI:214853 |
| KEYWORDS | transposon. |
| SEGMENT | 1 of 2 |
| SOURCE | X.laavis DNA. |
| ORGANISM | Xenopus laevis |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| | Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipoidea; |

| Db | 19 | TAGTACAGTTCTGCTACCTATTGCGATAGAGAACTTCAGGATTAATTCTCGGAAT | 78 |
|------------|--|---|----|
| QY | 76 | TT 77 | |
| Db | 79 | TT 80 | |
| RESULT 11 | | | |
| LOCUS | AF121764 | | |
| DEFINITION | AF121764 81 bp DNA linear INV 27-JUL-1999 Phytomonas staheli strain PLT4 trans-spliced leader SL-1, partial sequence. | | |
| ACCESSION | AF121764 | | |
| VERSION | AF121764.1 | GI:4927005 | |
| KEYWORDS | Phytomonas staheli. Phytomonas staheli Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Phytomonas. 1 (bases 1 to 81) Serrano,M.G., Campaner,M., Buck,G.A., Teixeira,M.M. and Camargo,E.P. PCR amplification of the spliced leader gene for the diagnosis of trypanosomatid parasites of plants and insects in methanol-fixed samples | | |
| REFERENCE | Serrano,M.G., Campaner,M., Buck,G.A., Teixeira,M.M. and Camargo,E.P. PCR amplification of the spliced leader gene for the diagnosis of trypanosomatid parasites of plants and insects in methanol-fixed samples | | |
| AUTHORS | Serrano,M.G., Campaner,M., Buck,G.A., Teixeira,M.M. and Camargo,E.P. Direct Submission Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| TITLE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| JOURNAL | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| MEDLINE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| PUBMED | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| REFERENCE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| AUTHORS | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| TITLE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| JOURNAL | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| MEDLINE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| PUBMED | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| REFERENCE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| AUTHORS | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| TITLE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| JOURNAL | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| MEDLINE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| PUBMED | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| REFERENCE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| AUTHORS | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| TITLE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| JOURNAL | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| MEDLINE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| PUBMED | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| REFERENCE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| AUTHORS | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| TITLE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| JOURNAL | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| MEDLINE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| PUBMED | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| REFERENCE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| AUTHORS | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| TITLE | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |
| JOURNAL | Submitted (21-JAN-1999) Parasitology, University of Sao Paulo, AV. Prof. Linneu Prestes, 1374, Sao Paulo, SP 05508-900, Brazil location/Qualifiers | | |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 61)
Suzuki,K., Miki,T., Kawamata,N., Hirotsawa,S., Yoshizawa,K., Kiyosawa,K. and Aoki,N.
Variant translocation of the BCL6 gene to immunoglobulin kappa light chain gene in B-cell lymphoma
Jpn. J. Cancer Res. 85 (9), 911-917 (1994)
JOURNAL 95049867
MEDLINE
REFERENCE 2 (bases 1 to 61)
AUTHORS Suzuki,K.
TITLE Direct Submission
JOURNAL Submitted (10-FEB-1994) Ken Suzuki, Tokyo Medical and Dental University, First department of Internal Medicine: 1-5-45 Yushima, Bunkyo-ku, Tokyo 113, Japan (Tel:03-3813-6111(ex.3217), Fax:03-3818-0448)
FEATURES
source
1. .61
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p11"
/cell_line="SUBL"
/cell_type="malignant B cell"
/tissue_type="lymphoma"
/germline
1. .61
/gene="Igk"
misc_feature
1. .61
/gene="Igk"
/product="immunoglobulin kappa light chain"
/note="joining region of chromosomes 2 and 3"
/db_xref="IMG/LOCID:D8525"
BASE COUNT 23 a 6 c 6 g 26 t
ORIGIN
Query Match 22.1%; Score 18.6; DB 9; Length 61;
Best Local Similarity 65.9%; Pred. No. 26+05;
Matches 27; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
OY 3 TTCATCTTCAGGTACAGACAGTGTGTCCTCGTTG 43
||||| ||||| | ||||| | ||||| |||||
Db 6 TTCATATTTCAAAATACACAGTTTCATGTATCTATTG 46
RESULT 13
AF454277 73 bp DNA linear VRT 09-MAY-2002
LOCUS Ficedula albicollis isolate Bc17 microsatellite z135 sequence.
DEFINITION AF454277
ACCESSION AF454277.1 GI:20514132
VERSION
KEYWORDS
SOURCE Ficedula albicollis.
ORGANISM Ficedula albicollis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauiria; Aves; Neognathae; Passeriformes; Muscicapidae; Ficedula.
1 (bases 1 to 73)
Primmer,C.R., Borge,T., Lindell,J. and Saetre,G.-P.
Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome
Mol. Ecol. 11 (3), 603-612 (2002)
JOURNAL 21918460
MEDLINE 11918793
PUBMED
REFERENCE 2 (bases 1 to 73)
AUTHORS Primmer,C.R., Borge,T. and Saetre,G.-P.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) Ecology and Systematics, University of Helsinki, Arkadiankatu 7, Helsinki 00014, Finland
FEATURES
source
1. .73
/organism="Ficedula albicollis"

/isolate="Bc17"
/db_xref="taxon:59894"
repeat_region
1. .73
/note="microsatellite z135; orthologous locus identified in Zosterops lateralis, Primmer et al. (2002)"
/rpt_type="tandem"
BASE COUNT 15 a 14 c 23 g 21 t
ORIGIN
Query Match 22.1%; Score 18.6; DB 5; Length 73;
Best Local Similarity 65.9%; Pred. No. 1.9e+05;
Matches 27; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
OY 10 TTCAGGTACAGACAGTGTGTCCTCGTGTGCTAA 50
||||| ||||| ||| ||||| ||||| |||||
Db 32 TTCAGCGCAGACAGCTTGTTGTGTGTGTCATGTGCTTA 72
RESULT 14
AB1641 83 bp DNA linear PAT 21-JAN-2000
LOCUS Sequence 18 from Patent WO9902671.
DEFINITION AB1641
ACCESSION AB1641
VERSION AB1641.1 GI:6731802
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 83)
AUTHORS Tawfik,D. and Griffiths,A.
TITLE IN VITRO SORTING METHOD
JOURNAL Patent: WO 9902671-A 18 21-JAN-1999;
MEDICAL RES COUNCIL (GB); TAWFIK DAN (GB)
FEATURES
source
1. .83
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 35 a 10 c 14 g 24 t
ORIGIN
Query Match 22.1%; Score 18.6; DB 6; Length 83;
Best Local Similarity 65.9%; Pred. No. 1.9e+05;
Matches 27; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
OY 41 TTGTGCTTAACCTTTAAGACGCGTGTTCACGAATTGATA 81
||| || ||||| ||||| | | | ||||| |
Db 21 TTTTGTTTAAGAGAGATATACATATGATTGAAA 61
RESULT 15
AX043859 61 bp DNA linear PAT 23-NOV-2000
LOCUS AX043859/c
DEFINITION Sequence 66 from Patent WO0063391.
ACCESSION AX043859
VERSION AX043859.1 GI:11342444
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 61)
AUTHORS Savidge,B., Lassner,M.W., Weiss,J.D. and Post-Beltemmler,D.
TITLE Nucleic acid sequences to proteins involved in tocopherol synthesis
JOURNAL Patent: WO 0063391-A 66 26-OCT-2000;
Calgene LLC (US)
FEATURES
source
1. .61
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"
BASE COUNT 20 a 12 c 16 g 13 t
ORIGIN
Query Match 21.9%; Score 18.4; DB 6; Length 61;

Wed Feb 19 10:48:50 2003

us-09-939-293-1_copy_56_139.lim.rge

Page 6

| | | | | | | |
|-----------------------|------------------|-----------|------------|------------|---------|----|
| Best Local Similarity | 69.48; | Pred. No. | 2.3e+05; | | | |
| Matches | 25; Conservative | 0; | Mismatches | 11; Indels | 0; Gaps | 0; |

```
Oy      41 TTGTGCTAACTTTAAGAAGCGGTGTTTCAGAAAT   76
        ||| | | | | | | | | | | | | | | |
Db      53 TTGTGCTTAACCAAGCACCAGTGTGTCTCAAAAT   18
```

Search completed: February 16, 2003, 04:26:47
Job time : 3189 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 03:13:16 : Search time 296 Seconds

(without alignments)
639,081 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84
Sequence: 1 actcattcttcaggtacag.....gtttctcagatgtataga 84

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2352344

Minimum DB seq length: 0

Maximum DB seq length: 84

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 20.8 | 24.8 | 60 | 24 | ABN45377 |
| 2 | 20.4 | 24.3 | 42 | 20 | AAA23305 |
| 3 | 20.4 | 24.3 | 80 | 22 | AA15381 |
| 4 | 20.4 | 24.3 | 81 | 22 | AA15380 |
| 5 | 20.2 | 24.0 | 50 | 18 | AA179460 |
| 6 | 19.8 | 23.6 | 80 | 16 | AA080399 |
| 7 | 19.4 | 23.1 | 83 | 22 | AA23612 |
| 8 | 19.2 | 22.9 | 65 | 15 | AA057728 |
| 9 | 18.8 | 22.4 | 60 | 24 | ABN36963 |

| | | | | | |
|----|------|------|----|----|----------|
| 10 | 18.6 | 22.1 | 51 | 22 | AA131832 |
| 11 | 18.6 | 22.1 | 66 | 22 | AA504820 |
| 12 | 18.6 | 22.1 | 83 | 20 | AA02762 |
| 13 | 18.4 | 21.9 | 42 | 20 | AA117569 |
| 14 | 18.4 | 21.9 | 60 | 24 | ABN36340 |
| 15 | 18.4 | 21.9 | 61 | 21 | AA061934 |
| 16 | 18.4 | 21.9 | 61 | 24 | AA062215 |
| 17 | 18.4 | 21.9 | 61 | 24 | AA596969 |
| 18 | 18.2 | 21.7 | 60 | 24 | ABN32614 |
| 19 | 18.2 | 21.7 | 60 | 24 | ABN40611 |
| 20 | 18.2 | 21.7 | 60 | 24 | AA056575 |
| 21 | 18.2 | 21.4 | 51 | 22 | AA056575 |
| 22 | 18.2 | 21.4 | 60 | 20 | AA056575 |
| 23 | 17.8 | 21.2 | 42 | 20 | AA117887 |
| 24 | 17.8 | 21.2 | 42 | 20 | AA091982 |
| 25 | 17.8 | 21.2 | 53 | 20 | AA035951 |
| 26 | 17.8 | 21.2 | 58 | 13 | AA034116 |
| 27 | 17.8 | 21.2 | 60 | 24 | ABN33955 |
| 28 | 17.6 | 21.0 | 45 | 21 | ABN96820 |
| 29 | 17.6 | 21.0 | 51 | 21 | AA026120 |
| 30 | 17.6 | 21.0 | 63 | 22 | AA088500 |
| 31 | 17.6 | 21.0 | 65 | 24 | ABN29697 |
| 32 | 17.6 | 21.0 | 65 | 24 | ABN30661 |
| 33 | 17.6 | 21.0 | 65 | 24 | ABN31939 |
| 34 | 17.6 | 21.0 | 65 | 24 | ABN53033 |
| 35 | 17.6 | 21.0 | 75 | 19 | AA020032 |
| 36 | 17.6 | 21.0 | 75 | 21 | AA071291 |
| 37 | 17.6 | 21.0 | 75 | 22 | AA031989 |
| 38 | 17.4 | 20.7 | 31 | 20 | AA034383 |
| 39 | 17.4 | 20.7 | 42 | 20 | AA017602 |
| 40 | 17.4 | 20.7 | 42 | 21 | AA064548 |
| 41 | 17.4 | 20.7 | 42 | 21 | AA064641 |
| 42 | 17.4 | 20.7 | 54 | 18 | AA063321 |
| 43 | 17.4 | 20.7 | 65 | 24 | ABN31237 |
| 44 | 17.4 | 20.7 | 69 | 22 | AA054772 |
| 45 | 17.4 | 20.7 | 75 | 24 | ABN54215 |

ALIGNMENTS

| | | |
|----------|---|-----------------------|
| RESULT 1 | ABN45377 | standard; DNA; 60 BP. |
| ID | ABN45377 | |
| AC | ABN45377 | |
| XX | | |
| DT | 15-JUL-2002 | (first entry) |
| DE | Human spliced transcript detection oligonucleotide seq ID NO:18125. | |
| XX | | |
| KW | Human; mouse; rat; splice transcript; detection; RNA transcript; | |
| KW | splice variant; transcriptome; oligonucleotide library; ss. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO200210449-A2. | |
| XX | | |
| PD | 07-FEB-2002. | |
| XX | | |
| PF | 20-JUL-2001; 2001MO-IB01903. | |
| XX | | |
| PR | 28-JUL-2000; 2000US-221607P. | |
| XX | | |
| PR | 02-MAY-2001; 2001US-287724P. | |
| XX | | |
| PA | (COMP-) COMPUGEN INC. | |
| XX | | |
| PI | Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S; | |
| XX | | |
| DR | WPI; 2002-257383/30. | |
| XX | | |
| PT | New oligonucleotide libraries comprising oligonucleotides which selectively hybridize to mRNAs transcribed from a transcription unit of | |

PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
XX
PS Example 1; SEQ ID 18125; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pot_sequences.
XX
SQ Sequence 60 BP; 14 A; 6 C; 21 G; 19 T; 0 other;
Query Match 24.8%; Score 20.8; DB 24; Length 60;
Best Local Similarity 64.6%; Pred. No. 8.2e+02;
Matches 31; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
OY 14 GGTACAGACAGTGTGTGTCCTGTCGCTACTTAAAGC 61
DB 13 GGTTCGAAAGGAAATGTTGTCCTTGTGCGCACAATTAAAGATGC 60
RESULT 2
AAA23305/c
ID AAA23305 standard; RNA; 42 BP.
AC AAA23305;
XX
XX 19-JUN-2000 (first entry)
DE Integrin subunit beta 3 hairpin ribozyme SEQ ID NO:6531.
XX
XX Human; aryl hydrocarbon nuclear transport; ARNT; TIE-2; angiogenesis;
KW Integrin alpha 6 subunit; integrin subunit beta 3; hairpin ribozyme;
KW hammerhead ribozyme; angiogenic factor; cytosolic; antidiabetic;
KW ophthalmologic; antiinflammatory; antiarthritic; antipsoriatic; ARMD;
KW dermatologic; RNA cleavage; cancer; diabetic retinopathy; arthritis;
KW age related macular degeneration; inflammation; neovascular glaucoma;
KW myopic degeneration; psoriasis; verruca vulgaris; angiofibroma;
KW tuberosus sclerosis; pot-wine stain; Sturge Weber syndrome;
KW Kippel-Trenaunay-Weber syndrome; Osler-Weber-Rendu syndrome; ss.
XX
XX Homo sapiens.
OS
XX
XX W09950403-A2.
PN
XX
XX 07-OCT-1999.
PD
XX
XX 24-MAR-1999; 99WO-US06507.
PF
XX
XX 27-MAR-1998; 98US-0079678.
PR
XX
XX (RIBO-) RIBOZYME PHARM INC.
PA
XX

PI Pavco PA, Roberts E, Jarvis T, Coeshott C, McSwigen JA;
XX
XX WPI; 1999-591315/50.
DR
XX
XX
PT Novel ribozymes for modulating the synthesis, expression and/or
PT stability of an mRNA encoding an angiogenic factors -
XX
XX Claim 14; Page 275; 305pp; English.
XX
XX The present invention describes enzymatic nucleic acid molecules with
CC RNA cleaving activity, which specifically cleave RNA encoded by an aryl
CC hydrocarbon nuclear transporter (ARNT) gene, an integrin subunit beta 3
CC gene, an integrin alpha 6 subunit gene, or a Tie-2 gene. AAA16775 to
CC AAA17167 and AAA17561 to AAA17622 represent ribozyme sequences for ARNT,
CC and AAA17168 to AAA17560 and AAA17623 to AAA17684 represent their
CC corresponding target sequences; AAA17685 to AAA18385 and AAA19087 to
CC AAA19154 represent ribozyme sequences for Tie-2, and AAA18386 to AAA19086
CC and AAA19155 to AAA19222 represent their corresponding target sequences;
CC AAA19223 to AAA20361 and AAA21501 to AAA21599 represent ribozyme
CC sequences for integrin alpha 6 subunit, and AAA20362 to AAA21500 and
CC AAA21596 to AAA21688 represent their corresponding target sequences;
CC AAA21689 to AAA22475 and AAA23263 to AAA23342 represent ribozyme sequence
CC for integrin subunit beta 3, and AAA22476 to AAA23262, AAA23343 to
CC AAA23422 represent their corresponding target sequences. The ribozymes of
CC the invention are used for modulating the synthesis, expression and/or
CC stability of an mRNA encoding angiogenic factor, especially ARNT,
CC integrin subunit beta-3, integrin subunit alpha-6, or Tie-2. They are
CC especially used to treat cancer, diabetic retinopathy, age related
CC macular degeneration (ARMD), inflammation, and arthritis, as well as
CC neovascular glaucoma, myopic degeneration, psoriasis, verruca vulgaris,
CC angiofibroma of tuberosus sclerosis, pot-wine stains, Sturge Weber
CC syndrome, Kippel-Trenaunay-Weber syndrome, Osler-Weber-Rendu syndrome,
CC and other syndromes and diseases related to the levels of ARNT, Tie-2,
CC integrin subunit alpha-6, or integrin subunit beta-3.
XX
SQ Sequence 42 BP; 15 A; 8 C; 10 G; 8 U; 1 other;
Query Match 24.3%; Score 20.4; DB 20; Length 42;
Best Local Similarity 69.2%; Pred. No. 1e+03;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
OY 12 CAGCTACAGACAGTGTGTGTCCTGTCGCTAA 50
DB 39 CAGTATAGTACNTGTTCTGTCGCTCACTTGGCTAA 1
RESULT 3
AAD15381/c
ID AAD15381 standard; DNA; 80 BP.
AC AAD15381;
XX
XX 15-NOV-2001 (first entry)
DE Oligo 10R, to construct pDual prokaryotic/mammalian expression vector.
KW Dual expression vector; gene expression; prokaryotic cell;
KW eukaryotic cell; ss.
XX
XX Synthetic.
OS
XX
XX US2001016351-A1.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 31-OCT-1997; 97US-0961888.
PF
XX
XX 31-OCT-1997; 97US-0961888.
PR
XX
XX (SORG/) SORGE J A.
PA (PADG/) PADGETT K A.
XX
XX Sorge JA, Padgett KA;
PI

```
XX
DR WPI; 2001-496443/54.
XX
PT New dual expression vector, useful for high level protein expression in
PT prokaryotic and eukaryotic cells, comprises cloning site, transcription
PT elements and replication elements -
XX
PS Example 1; Page 6; 28bp; English.
XX
CC The invention relates to dual expression vectors useful for high level
CC expression of any gene sequence in both prokaryotic and eukaryotic cells.
CC This vector comprises a cloning site, transcription elements sufficient
CC to permit transcription of a polynucleotide in both prokaryotic and
CC eukaryotic cell and replication elements sufficient to permit the
CC replication of the vector. The vector also facilitates the isolation,
CC manipulation and study of eukaryotic genes, gene fragments and cDNA
CC transcripts. The present sequence is a synthetic oligonucleotide used
CC to construct pnal prokaryotic/mammalian expression vector.
XX
SQ Sequence 80 BP; 22 A; 18 C; 15 G; 25 T; 0 other;
Query Match 24.3%; Score 20.4; DB 22; Length 80;
Best Local Similarity 58.1%; Pred. No. 1.2e+03;
Matches 36; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 4 TCATTTCTTCAGTACAGACAGTGTGTGTCCTGTCGTGCGCTACTTTAGAACGG 63
DB 71 TCACATATAGGGGAATTGTGAGCGGATACAAATTCCTTTGTTTACTTTAGAACGAG 12
QY 64 TG 65
DB 11 AG 10
Db
RESULT 4
AADI5380
ID AADI5380 standard; DNA; 81 BP.
XX
AC AADI5380;
XX
DT 15-NOV-2001 (first entry)
XX
DE Oligo 10F, to construct pDual prokaryotic/mammalian expression vector.
XX
KM Dual expression vector; gene expression; prokaryotic cell;
XX eukaryotic cell; ss.
XX
OS Synthetic.
XX
PN US2001016351-A1.
XX
PD 23-AUG-2001.
XX
PF 31-OCT-1997; 97US-0961888.
XX
PR 31-OCT-1997; 97US-0961888.
XX
PA (SORG/) SORGE J A.
XX (PADG/) PADGETT K A.
XX
PI SORGE JA, PADGETT KA;
XX
DR WPI; 2001-496443/54.
XX
PT New dual expression vector, useful for high level protein expression in
PT prokaryotic and eukaryotic cells, comprises cloning site, transcription
PT elements and replication elements -
XX
PS Example 1; Page 6; 28bp; English.
XX
CC The invention relates to dual expression vectors useful for high level
CC expression of any gene sequence in both prokaryotic and eukaryotic cells.
CC This vector comprises a cloning site, transcription elements sufficient
```

```
CC to permit transcription of a polynucleotide in both prokaryotic and
CC eukaryotic cell and replication elements sufficient to permit the
CC replication of the vector. The vector also facilitates the isolation,
CC manipulation and study of eukaryotic genes, gene fragments and cDNA
CC transcripts. The present sequence is a synthetic oligonucleotide used
CC to construct pnal prokaryotic/mammalian expression vector.
XX
SQ Sequence 81 BP; 25 A; 16 C; 18 G; 22 T; 0 other;
Query Match 24.3%; Score 20.4; DB 22; Length 81;
Best Local Similarity 58.1%; Pred. No. 1.2e+03;
Matches 36; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 4 TCATTTCTTCAGTACAGACAGTGTGTGTCCTGTCGTGCGCTACTTTAGAACGG 63
DB 15 TCACATATAGGGGAATTGTGAGCGGATACAAATTCCTTTGTTTACTTTAGAACGAG 74
QY 64 TG 65
DB 75 AG 76
Db
RESULT 5
AAV79460/C
ID AAV79460 standard; DNA; 50 BP.
XX
AC AAV79460;
XX
DT 16-MAR-1999 (first entry)
XX
DE Staphylococcus aureus contig SEQ ID #5149.
XX
KM Computer readable medium; vaccine; S.aureus infection; immunodetection;
KM cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KM skin infection; surgical wound infection; scalded skin syndrome;
KM toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
PN EP786519-A2.
XX
PD 30-JUL-1997.
XX
PF 07-JAN-1997; 97EP-0100117.
XX
PR 05-JAN-1996; 96US-0009861.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
XX Rosen CA;
XX
DR WPI; 1997-374922/35.
XX
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX
PS Claim 1; Page 3177; 3271pp; English.
XX
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
```

CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.

XX Sequence 50 BP; 23 A; 11 C; 10 G; 5 T; 1 other;

Query Match 24.0%; Score 20.2; DB 18; Length 50;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 4 TCATTCTTCAGGTACACAGTGTGTGTTCTGTGTG 45
DB 46 TCTTCTCCGGTACTTAAGATGTTCAGTTCGGGTGTG 5

RESULT 6
AA080399/c
ID AA080399 standard; cDNA; 80 BP.

XX AA080399;

DT 17-JUL-1995 (first entry)

XX CDR-grafted L243-gH primer R4897.

XX Humanized antibody; antibody engineering; MHC class I;
KW major histocompatibility region; HLA; monoclonal antibody; MAb;
KW L243-gH; immunological disease; transplantation; light chain;
KW heavy chain; variable region; complementarity determining region;
KW CDR; primer; polymerase chain reaction; PCR; ss.

XX Synthetic.

OS WO9429451-A.

XX 22-DEC-1994.

XX 15-JUN-1994; 94WO-GB01291.

XX 16-JUN-1993; 93GB-0012415.

XX 27-JAN-1994; 94GB-0001597.

XX 09-FEB-1994; 94GB-0002499.

XX 29-MAR-1994; 94GB-0006222.

XX (CLUT) CELUTECH LTD.

XX Athwal DS, Bodmer MW, Emlage JS, Morgan SA;

XX WPI; 1995-036480/05.

XX New humanised anti-HLA DR antibodies - used for diagnosis and

XX treatment of immunological diseases and transplantation related

XX conditions

XX Disclosure; Page 35; 68pp; English.

XX L243 is a mouse MAb raised against human MHC class II. A

XX CDR-grafted humanized antibody heavy chain, L243-gH, has

XX a variable region domain comprising acceptor frameworks derived

XX from human gr. consensus sequence and L243 donor antigen binding

XX regions. L243-gH was assembled from overlapping oligonucleotides

XX using the PCR primers given in AA080393-403. Products were expressed

XX in Escherichia coli LM1035.

XX Sequence 80 BP; 22 A; 26 C; 20 G; 12 T; 0 other;

Query Match 23.6%; Score 19.8; DB 16; Length 80;
Best Local Similarity 69.2%; Pred. No. 2e+03;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 28 TTGTGTCTCTGTGTGCTAAGTAAAGCGGTGT 66

DB 48 TTTGGGTGTCTCTCTCTGTGTGACACAGAGT 10

RESULT 7
AAS23612
ID AAS23612 standard; DNA; 83 BP.

XX AAS23612;

DT 04-DEC-2001 (first entry)

XX Tetracycline promoter upstream PCR primer (Tet-Up) #48.

XX Gene identification; essential gene; GRACE; pathogenic fungus;

XX gene replacement and conditional expression; fungal infection;

XX PCR primer; Tet-Up; tetracycline promoter; ss.

XX Escherichia coli.

XX Candida albicans.

XX WO200160975-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05551.

XX 18-FEB-2000; 2000US-0183534.

XX (ELIT-) ELITRA PHARM INC.

XX Roemer T, Jiang B, Boone C, Bussey H;

XX WPI; 2001-489080/53.

XX Identifying genes essential to fungal metabolisms and identifying

XX potential therapeutic agents that target these genes -

XX Disclosure; Page 285; 324pp; English.

XX The present invention relates to novel methods for constructing fungal

XX strains useful for identification and validation of gene products as

XX targets for therapeutic agents, for creating a collection of identified

XX essential genes, and screening assays for the discovery of new drugs.

XX The invention provides the GRACE (gene replacement and conditional

XX expression) method for the construction of mutant organisms referred to

XX as GRACE strains of the organism. The invention can be applied to any

XX organism, particularly a pathogenic fungus e.g. Candida albicans,

XX Aspergillus fumigatus and Cryptococcus neoformans. The methods are

XX useful to identify agents that may be used in the treatment of fungal

XX infections. AAS23565-AAS23625 represent tetracycline promoter upstream

XX PCR primers (Tet-Up) used in the methods of the present invention.

XX Sequence 83 BP; 20 A; 5 C; 21 G; 37 T; 0 other;

Query Match 23.1%; Score 19.4; DB 22; Length 83;
Best Local Similarity 55.1%; Pred. No. 2.7e+03;
Matches 38; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

OY 8 TCTTCAGGTACACAGTGTGTGTTCTGTGTGCTAAGTAAAGCGGTGT 67
DB 4 TATCTTTTAAATGATGTGTGTGTGTGTGATATATATTTGGAAGAAATGAT 63
OY 68 TCTCAGAAAT 76
DB 64 TGTGATTAT 72

RESULT 8
AA057728/c
ID AA057728 standard; DNA; 65 BP.
XX AA057728;


```

XX 12-AUG-1994 (first entry)
DT
XX
XX
DE Viral immunosuppressive region-related clone I-3 from human brain.
XX
XX Immunosuppression; retrovirus; transmembrane protein; cancer;
KW conserved sequence; autoimmune disease; graft vs. host disease;
XX mammalian type C virus; polymerase chain reaction; human; brain; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH misc_feature 3..65
FT /tag= a
FT /note= "encodes amino acid sequence corresp. to
FT Immunosuppressive region in IV-proteins
FT of mammalian type C viruses"
XX
XX WO9403493-A.
PN
XX 17-FEB-1994.
PD
XX
XX 08-AUG-1993; 93WO-SE00671.
PF
XX
XX 10-AUG-1992; 92SE-0002318.
PR
XX
XX (REPL-) REPLICO MEDICAL AB.
PA
XX Blomberg J, Lindeskog M, Pipkorn R;
PI
XX MPI; 1994-065608/08.
DR
XX P-PSDB; AAR46656.
XX
XX Peptide(s) derived from immunosuppressive region in retro-viruses
PT - used in diagnosis, treatment and prevention of retrovirus
PT disease, auto-immune diseases and malignant diseases
PT
XX
XX Disclosure: Page 32; 41pp; English.
XX
XX Evolutionarily conserved sequences corresponding to an
CC immunosuppressive region in retroviral transmembrane proteins were
CC amplified by PCR from human brain DNA and reverse transcribed RNA.
CC The primers used were CDTML1 and CDTMR2 (AA057718-9). Amplified
CC sequences could be classified into 3 groups (I, II and III).
CC
XX AA057728 is a member of Group I.
XX
XX Sequence 65 BP; 22 A; 14 C; 14 G; 15 T; 0 other;
SQ
XX
XX Query Match 22.9%; Score 19.2; DB 15; Length 65;
XX Best Local Similarity 56.2%; Pred. No. 3e+03; Indels 0; Gaps 0;
XX Matches 36; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
XX
OY 4 TCATTCTTCAGGTACAGACAGTGTGTGCTGCTGTTGGCTAACTTTAAGACGG 63
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 64 TCTTCATTAAAGAAATATACAGAGTCTCTCTTTTTCAGCTGTAGTAAGTCAAGCCTCG 5
OY 64 TGT 67
DB |||
DB 4 CGGT 1

```

```

XX OS Homo sapiens.
XX
XX WO200210449-A2.
PN
XX
XX 07-FEB-2002.
PD
XX
XX 20-JUL-2001; 2001WO-1B01903.
PF
XX
XX 28-JUL-2000; 2000US-221607P.
PR
XX 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
PA
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI
XX MPI; 2002-257383/30.
DR
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
XX Example 1; SEQ ID 9711; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptsomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABR27253 to ABR59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pt_sequences.
XX
XX Sequence 60 BP; 22 A; 16 C; 9 G; 13 T; 0 other;
SQ
XX
XX Query Match 22.4%; Score 18.8; DB 24; Length 60;
XX Best Local Similarity 68.4%; Pred. No. 4.1e+03;
XX Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
XX
OY 23 AGTGTGTGTGTCTCCTGTTGGCTAACTTTAAGAG 60
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 58 ACTCTGTGGGTTTCCAGTGTGATTAATTTAAGCAG 21

```

```

RESULT 9
ABN36963/C
ID ABN36963 standard; DNA; 60 BP.
XX
XX ABN36963;
AC
XX
XX 15-JUL-2002 (first entry)
DT
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:9711.
DE
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.

```

```

RESULT 10
AAL31832
ID AAL31832 standard; DNA; 51 BP.
XX
XX AAL31832;
AC
XX
XX 24-JAN-2002 (first entry)
DT
XX
XX Human SNP oligonucleotide #5040.
DE
XX
XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;

```

KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinase; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease; ss.
 XX
 OS Homo sapiens.
 PN WO200147944-A2.
 XX
 XX 05-JUL-2001.
 PD
 XX 28-DEC-2000; 2000WO-US35498.
 PF
 XX 28-DEC-1999; 99US-0173419.
 PR 27-DEC-2000; 2000US-0173419.
 XX
 PA (CORA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 XX WPI: 2001-465210/50.
 DR
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -
 XX
 PS Claim 1; Page 2836; 4143pp; English.
 XX
 CC The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kinases, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.
 CC
 XX Sequence 51 BP; 12 A; 8 C; 6 G; 25 T; 0 other;
 SO
 Query Match 22.1%; Score 18.6; DB 22; Length 51;
 Best Local Similarity 72.7%; Pred. No. 4.6e+03;
 Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 OY 1 ACTTCATTCTTCAGGTACGACAGTGTTCGT 33
 Db 18 ACTTCATTCTTCAGGTACGACAGTGTTCGT 50
 RESULT 11
 ID AAS04820/c
 AC AAS04820; standard; DNA; 66 BP.
 XX
 AC AAS04820;
 XX
 DT 07-SEP-2001 (first entry)
 XX
 XX Synthetic gene shGLP/RA, PCR primer #2.
 DE
 KW shGLP/RA: PCR primer; ss; immunogenic epitope cluster; IEC;
 KW synthetic human chondrocyte glycoprotein 69/cartilage link protein;
 KW autoantigen; autoimmune disease; multiple sclerosis;
 KW rheumatoid arthritis; insulin-dependent diabetes mellitus;
 KW myasthenia gravis; uveitis; autoimmune hepatitis; thyroiditis; insulinitis;

KW orchitis; idiopathic thrombocytopenic purpura; inflammatory disease;
 KW Crohn's disease; ulcerative colitis.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO200131037-A2.
 XX
 PD 03-MAY-2001.
 PD
 XX 26-OCT-2000; 2000WO-IL00688.
 PE
 XX 27-OCT-1999; 99IT-0132611.
 PR
 XX (YEDA) YEDA RES & DEV CO LTD.
 PA
 XX Ben-Nun A, Kerlero De Rosbo N, Sappier GP;
 PI
 DR WPI: 2001-300515/31.
 XX
 XX Novel synthetic human target autoantigen gene useful for treating
 PT autoimmune diseases such as multiple sclerosis, insulin-dependent
 PT diabetes mellitus, rheumatoid arthritis, myasthenia gravis, and uveitis
 PT
 XX
 PS Example 7; Fig 50; 182pp; English.
 XX
 CC The sequence represents a PCR primer used in the construction of
 CC a nucleic acid encoding shGLP/RA (synthetic human chondrocyte
 CC glycoprotein 69/cartilage link protein) containing immunogenic epitope
 CC clusters (IEC) from chondrocyte glycoprotein 69 and cartilage link
 CC protein. The synthetic human target autoantigen genes of the invention
 CC comprise sequences coding for at least 2 IECs of autoantigen(s) related
 CC to a specific autoimmune disease. The synthetic human target autoantigen
 CC genes are useful for treating autoimmune diseases such as multiple
 CC sclerosis, insulin-dependent diabetes mellitus, rheumatoid arthritis,
 CC myasthenia gravis, uveitis, autoimmune hepatitis, thyroiditis, insulinitis,
 CC Crohn's disease, idiopathic thrombocytopenic purpura, and inflammatory diseases
 CC (Crohn's disease, ulcerative colitis). The synthetic human target
 CC autoantigen genes are also useful for diagnosis and/or monitoring the
 CC progression of the autoimmune disease.
 CC
 XX Sequence 66 BP; 16 A; 21 C; 14 G; 15 T; 0 other;
 SO
 Query Match 22.1%; Score 18.6; DB 22; Length 66;
 Best Local Similarity 65.9%; Pred. No. 4.9e+03;
 Matches 27; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 OY 41 TTGTGCTTAACCTTAAGAGCGGTGTTCTCAGAAATGATA 81
 Db 49 TGGAGCTATCTTACGAAAGCGGTGAGACCAAGTCTCA 9
 RESULT 12
 ID AAX02762
 AC AAX02762; standard; DNA; 83 BP.
 XX
 AC AAX02762;
 XX
 DT 14-MAY-1999 (first entry)
 XX
 DE Oligonucleotide #18.
 DE
 KW Primer; genetic element; compartmentalised; microcapsule; screening;
 KW modulator; isolation; ss.
 XX
 OS Synthetic.
 OS
 PN WO9902671-A1.
 XX
 XX 21-JAN-1999.
 PD
 XX 29-JUN-1998; 98WO-GB01889.

```
XX 25-MAR-1998: 98GB-0006393.
PR 07-JUL-1997: 97GB-0014300.
XX
XX (MEDI-) MEDICAL RES COUNCIL.
XX
XX Griffiths A, Tawfik D;
XX
XX WPI: 1999-120872/10.
XX
XX Isolating genetic elements encoding desired gene products from
PT microcapsules - useful for in vitro screening for compounds that
PT modulate the activity of gene products
XX
XX Disclosure: Page 85; 101pp; English.
XX
XX This invention describes a novel method for isolating genetic elements
XX encoding desired gene products. The method comprises expressing genetic
XX elements that are compartmentalized in microcapsules, and sorting the
XX genetic elements, which produce the desired gene products. The genetic
XX elements are useful for screening for compounds that modulate the
XX activity of a gene product. Polypeptide gene products are useful for
XX preparing compounds, comprising desired gene products. The method
XX provides an in vitro way of providing both nucleic acids and proteins
XX having a full range of biochemical and biological activities, and
XX combines several processes to isolate a desired product or activity.
XX
XX Sequence 83 BP; 35 A; 10 C; 14 G; 24 T; 0 other:
XX
XX Query Match 22.1%; Score 18.6; DB 20; Length 83;
XX Best Local Similarity 65.9%; Pred. No. 5.1e+03;
XX Matches 27; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
XX
OY 41 TTCTGCTAAGCTTTAGAGCGGTCTTCTCAGATTTGATA 81
XX ||||| ||||||| ||||| ||||| ||||| |||||
DB 21 TTTTGTTAAGCTTTAGAGAGATATACATATGATTGAAA 61
XX
RESULT 13
XX AAA17569/c
XX ID AAA17569 standard; RNA; 42 BP.
XX
XX AAA17569;
XX
XX 19-JUN-2000 (first entry)
XX
XX Aryl hydrocarbon nuclear transport hairpin ribozyme SEQ ID NO:795.
XX
XX Human; aryl hydrocarbon nuclear transport; ARNT; Tie-2; angiogenesis;
XX Integrin alpha 6 subunit; Integrin subunit beta 3; hairpin ribozyme;
XX hammerhead ribozyme; angiogenic factor; cytosolic; antidiabetic;
XX ophthalmologic; antiinflammatory; antiarthritic; antipsoriatic; ARMD;
XX dermatologic; RNA cleavage; cancer; diabetic retinopathy; arthritis;
XX age related macular degeneration; inflammation; neovascular glaucoma;
XX myopic degeneration; psoriasis; verruca vulgaris; angiodiroma;
XX tuberosus sclerosis; pot-wine stain; Sturge Weber syndrome;
XX Kippel-Trenaunay-Weber syndrome; Osler-Weber-Rendu syndrome; ss.
XX
XX Homo sapiens.
XX
XX WO9950403-A2.
XX
XX 07-OCT-1999.
XX
XX 24-MAR-1999; 99WO-US06507.
XX
XX 27-MAR-1998; 98US-0079678.
XX
XX (RIBO-) RIBOZYME PHARM INC.
XX
XX Payco PA, Roberts E, Jarvis T, Coesholt C, McSwiggen JA;
XX
XX WPI: 1999-591315/50.
XX
```

```
XX Novel ribozymes for modulating the synthesis, expression and/or
PT stability of an mRNA encoding an angiogenic factors
PT
XX Claim 14; Page 88; 305pp; English.
XX
XX The present invention describes enzymatic cleave acid molecules with
XX RNA cleaving activity, which specifically cleave RNA encoded by an aryl
XX hydrocarbon nuclear transporter (ARNT) gene, an Integrin subunit beta 3
XX gene, an Integrin alpha 6 subunit gene, or a Tie-2 gene. AAA16775 to
XX AAA17167 and AAA17561 to AAA17622 represent ribozyme sequences for ARNT,
XX and AAA17168 to AAA17560 and AAA17623 to AAA17684 represent their
XX corresponding target sequences; AAA17685 to AAA18385 and AAA19087 to
XX AAA19154 represent ribozyme sequences for Tie-2, and AAA18386 to AAA19086
XX and AAA19155 to AAA19222 represent their corresponding target sequences;
XX AAA19223 to AAA20361 and AAA21501 to AAA21595 represent ribozyme
XX sequences for Integrin alpha 6 subunit, and AAA20362 to AAA21500 and
XX AAA21596 to AAA21688 represent their corresponding target sequences;
XX AAA21689 to AAA22475 and AAA23263 to AAA23342 represent ribozyme sequence
XX for Integrin subunit beta 3, and AAA22476 to AAA23262, AAA23343 to
XX AAA23422 represent their corresponding target sequences. The ribozymes of
XX the invention are used for modulating the synthesis, expression and/or
XX stability of an mRNA encoding angiogenic factor, especially ARNT,
XX Integrin subunit beta-3, Integrin subunit alpha-6, or Tie-2. They are
XX especially used to treat cancer, diabetic retinopathy, age related
XX macular degeneration (ARMD), inflammation, and arthritis, as well as
XX neovascular glaucoma, myopic degeneration, psoriasis, verruca vulgaris,
XX angiodiroma of tuberosus sclerosis, pot-wine stains, Sturge Weber
XX syndrome, Kippel-Trenaunay-Weber syndrome, Osler-Weber-Rendu syndrome,
XX and other syndromes and diseases related to the levels of ARNT, Tie-2,
XX Integrin subunit alpha-6, or Integrin subunit beta-3.
XX
XX Sequence 42 BP; 15 A; 8 C; 12 G; 6 U; 1 other:
XX
XX Query Match 21.9%; Score 18.4; DB 20; Length 42;
XX Best Local Similarity 67.6%; Pred. No. 5.2e+03;
XX Matches 25; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
XX
OY 12 CAGGTACAGACAGTCTTGCTGCTCTCTGCTCGCT 48
XX ||||| | ||||| ||||| ||||| |||||
DB 39 CAGGTATGTCACNTGTTCTCTGCTCCACTTCGCT 3
XX
RESULT 14
XX ABN36340
XX ID ABN36340 standard; DNA; 60 BP.
XX
XX ABN36340;
XX
XX 15-JUL-2002 (first entry)
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:9086.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
XX
XX WO200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-IB01903.
XX
XX 28-JUL-2000; 2000US-221607P.
XX
XX 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX WPI: 2002-257383/30.
XX
```

PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -

PS Example 1; SEQ ID 9088; 47pp; English.
 YY

CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition, to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN7253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from Wipo
CC at [ftp.wipo.int/pub/published_pctl_sequences](http://wipo.int/pub/published_pctl_sequences).

Sequence 60 BP; 22 A; 6 C; 21 G; 11 T; 0 other;

| | | | | |
|-----------------------|--------------|--------------------|---------------|------------|
| Query Match | 21.9%; | Score 18.4; | DB 24; | Length 60; |
| Best Local Similarity | 78.6%; | Pred. No. 5.6e+03; | | |
| Matches 22; | Conservative | 0; | Mismatches 6; | Indels 0; |
| | | | | Gaps 0 |

QY 55 AAGAAGCGGTGTTCTCAGAAATTGATAA 82
| | | | |
Db 15 AGGAACGGGCTTCTCTGAATTTGGTAA 42

RESULT 15
AAC61934/c
ID AAC61934 standard; DNA; 61 BP.
...

AAC61934 ;

DT 06-MAR-2001 (first entry)

PCR primer for DNA encoding a prenyltransferase SLR0926 5' region.

KM Preylnylsferase; ATPp1; ATPp2; ATPp3; ATPp4; ATPp5; ATPp6; ATPp7;
 KM ATPp8; ATPp9; ATPp10; ATPp12; tocopherol; homogentistic acid;
 KM phyllypyrophosphate; 2-methyl-6-phytylbenzoquinol; antioxidant;
 KM nutritional supplement; PCR primer; ss.

Synechocystis sp.

PN WO200063391-A2.

PD 26-OCT-2000.

14-APR-2000; 2000WO-US10368.

15-APR-1999; 99US-0129899.
30-JUL-1999; 99US-0146461.

(CALJ) CALGENE LLC.

Savidge B, Lassner MW, Weiss JD, Post-Beitlenmiller D;

DR WPI; 2000-647519/62

PT An isolated nucleic acid sequence encoding prenyltransferase used to transform plant cells to increase the production of tocopherols -

PS Example 4; Page 29; 114pp; English.
xy

PCR primers AAC61933-34 were used to amplify prenyltransferase. The specification describes prenyltransferases designated *ATP11*, *ATP12*, *ATP13*, *ATP14*, *ATP15*, *ATP16*, *ATP17*, *ATP18*, *ATP19*, *ATP10*, *ATP11*, and *ATP12*. The biosynthesis of alpha-tocopherol in higher plants involves the condensation of homogentisic acid and phytylpyrophosphate to form 2-methyl-6phytylbenzoquinol, which can form various tocopherols. The prenyltransferase polynucleotides are useful in transforming host cells to alter the expression of prenyltransferase in these cells. The transformed cells are used in the production of tocopherols which are of use in the pharmaceutical industry as antioxidants and also in the food industry as nutritional supplements.

50 Sequence 61 BP; 20 A; 12 C; 16 G; 13 T; 0 other;

| | | | | |
|-----------------------|-------|-------------------|-------|---------------|
| Query Match | 21.9% | Score 18.4 | DB 21 | Length 61 |
| Best Local Similarity | 69.4% | Pred. No. 5.6e+03 | | |
| Matches | 25 | Conservative | 0 | Mismatches 11 |
| | | | | Indels 0 |
| | | | | Gaps 0 |

Oy 41 TTGTGGCTACTTTAAGAACGCGTGTTTCTCAGAAT 76
 ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 TTTGTGCTTACCCAAGCACGTTGTGTCCTCAAAAT 180

Search completed: February 16, 2003, 03:33:26
Job time : 297 secs

Page 8

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 03:28:12 ; Search time 67 Seconds
(without alignments)
384,490 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84
Sequence: 1 actcattcttcaggtacag.....gttcctcagattgataaga 84

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 672730

Minimum DB seq length: 0
Maximum DB seq length: 84

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--|
| C 1 | 19.8 | 23.6 | 80 | 4 | US-08-569-147-68 Sequence 68, Appl |
| C 2 | 18 | 21.4 | 39 | 3 | US-08-907-598-4 Sequence 4, Appl |
| C 3 | 18 | 21.4 | 80 | 4 | US-09-453-702B-168 Sequence 168, Appl |
| C 4 | 17.8 | 21.2 | 47 | 4 | US-09-641-638-982 Sequence 982, Appl |
| C 5 | 17.6 | 21.0 | 75 | 3 | US-09-023-082A-122 Sequence 122, Appl |
| C 6 | 17.4 | 20.7 | 31 | 4 | US-09-270-140A-45 Sequence 45, Appl |
| C 7 | 17.4 | 20.7 | 54 | 4 | US-08-679-645-1196 Sequence 1196, Appl |
| C 8 | 17.2 | 20.5 | 84 | 3 | US-08-976-413A-317 Sequence 317, Appl |
| C 9 | 17 | 20.2 | 39 | 4 | US-09-387-300-33 Sequence 33, Appl |
| C 10 | 17 | 20.2 | 42 | 2 | US-08-805-918-33 Sequence 33, Appl |
| C 11 | 17 | 20.2 | 42 | 2 | US-08-811-028-21 Sequence 21, Appl |
| C 12 | 16.8 | 20.0 | 49 | 1 | US-08-222-177A-361 Sequence 361, Appl |
| C 13 | 16.8 | 20.0 | 69 | 4 | US-09-486-356-3 Sequence 3, Appl |
| C 14 | 16.8 | 20.0 | 76 | 3 | US-09-023-082A-99 Sequence 99, Appl |
| C 15 | 16.8 | 20.0 | 76 | 4 | US-09-218-44A-20 Sequence 20, Appl |
| C 16 | 16.6 | 19.8 | 38 | 1 | US-07-971-160-43 Sequence 43, Appl |
| C 17 | 16.6 | 19.8 | 38 | 1 | US-08-336-241-43 Sequence 43, Appl |
| C 18 | 16.6 | 19.8 | 38 | 2 | US-08-465-273-43 Sequence 43, Appl |
| C 19 | 16.6 | 19.8 | 38 | 2 | US-09-119-024-43 Sequence 43, Appl |
| C 20 | 16.6 | 19.8 | 38 | 2 | US-08-417-226-43 Sequence 43, Appl |
| C 21 | 16.6 | 19.8 | 38 | 4 | US-09-196-131-43 Sequence 43, Appl |
| C 22 | 16.6 | 19.8 | 38 | 4 | US-08-643-732-43 Sequence 43, Appl |
| C 23 | 16.6 | 19.8 | 55 | 4 | US-09-620-958A-1 Sequence 1, Appl |
| C 24 | 16.6 | 19.8 | 62 | 1 | US-07-971-819A-37 Sequence 37, Appl |
| C 25 | 16.6 | 19.8 | 62 | 1 | US-08-475-231-37 Sequence 37, Appl |
| C 26 | 16.6 | 19.8 | 66 | 4 | US-08-591-632-64 Sequence 64, Appl |
| C 27 | 16.6 | 19.8 | 66 | 4 | US-09-611-451-64 Sequence 64, Appl |

| | | | | | | |
|------|------|------|----|---|---------------------|---------------------|
| 28 | 16.6 | 19.8 | 72 | 5 | PCR-US92-01015-3 | Sequence 3, Appl |
| 29 | 16.6 | 19.8 | 84 | 3 | US-09-023-082A-126 | Sequence 126, Appl |
| 30 | 16.6 | 19.8 | 84 | 3 | US-09-023-082A-130 | Sequence 130, Appl |
| C 31 | 16.4 | 19.5 | 18 | 4 | US-08-585-684B-2723 | Sequence 2723, Appl |
| C 32 | 16.4 | 19.5 | 18 | 4 | US-09-038-073-2723 | Sequence 2723, Appl |
| C 33 | 16.4 | 19.5 | 77 | 1 | US-08-384-708A-194 | Sequence 194, Appl |
| C 34 | 16.4 | 19.5 | 77 | 4 | US-08-687-421-286 | Sequence 286, Appl |
| C 35 | 16.4 | 19.5 | 80 | 1 | US-08-471-985A-88 | Sequence 88, Appl |
| C 36 | 16.4 | 19.5 | 80 | 5 | PCR-US95-12401A-88 | Sequence 88, Appl |
| C 37 | 16.2 | 19.3 | 47 | 4 | US-09-338-907-306 | Sequence 306, Appl |
| C 38 | 16.2 | 19.3 | 47 | 4 | US-09-218-207-306 | Sequence 306, Appl |
| C 39 | 16.2 | 19.3 | 52 | 1 | US-08-222-177A-418 | Sequence 418, Appl |
| C 40 | 16.2 | 19.3 | 62 | 1 | US-08-322-177A-56 | Sequence 56, Appl |
| C 41 | 16.2 | 19.3 | 79 | 1 | US-08-384-708A-211 | Sequence 211, Appl |
| C 42 | 16.2 | 19.3 | 79 | 2 | US-08-470-939-9 | Sequence 9, Appl |
| C 43 | 16.2 | 19.3 | 79 | 4 | US-08-687-421-303 | Sequence 303, Appl |
| C 44 | 16.2 | 19.3 | 79 | 5 | PCR-US96-09452-9 | Sequence 9, Appl |
| C 45 | 16 | 19.0 | 30 | 1 | US-08-050-319B-15 | Sequence 15, Appl |

ALIGNMENTS

```
RESULT 1
US-08-569-147-68/C
; Sequence 68, Application US/08569147
; Patent No. 6180377
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMANISED ANTIBODIES
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 618037719, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/569,147
; FILING DATE: 25-March-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yaliko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CARP-0047
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-569-147-68

Query Match 23.6%; Score 19.8; DB 4; Length 80;
Best Local Similarity 69.2%; Pred. No. 1.2e+02;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
```

```
QY 28 TTGTGTCTCTCTGTGTGCTAAGACGGGTGT 66
    ||| ||||| ||||| ||| ||| ||| |||
Db 48 TTGGGTCTCTCTCTCTCTCTGTCGACACAGAGT 10
```

```

RESULT 2
US-08-907-598-4/c
: Sequence 4, Application US/08907598
: Patent No. 6139833
: GENERAL INFORMATION:
: APPLICANT: Burgess, Rob
: APPLICANT: Friedrich, Glenn
: APPLICANT: Zambrowicz, Brian
: APPLICANT: Sands, Arthur
: TITLE OF INVENTION: TARGETED GENE DISCOVERY
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: PENNIE & EDMONDS LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/907,598
: FILING DATE:
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 8535-015-999
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 493-4935
: TELEFAX: (650) 493-5556
: TELEX: 661141
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 39 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: unknown
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ANTI-SENSE: NO
US-08-907-598-4

Query Match 21.4%; Score 18; DB 3; Length 39;
Best Local Similarity 70.6%; Pred. No. 4.4e+02;
Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 27 TTTGTCGTCTCTGCTGTGGCTACTTTTAAGAG 60
      1 1111 11111111 1 1 1 1 1 1 1 1
Db 39 TGTGTAGTTCCTGTCGTGACTCAATTGGTGTG 6
      1 1111 11111111 1 1 1 1 1 1 1 1

RESULT 3
US-09-453-702B-168/c
: Sequence 168, Application US/09453702B
: Patent No. 6365723
: GENERAL INFORMATION:
: APPLICANT: Blattner, Frederick R.
: APPLICANT: Burland, Valerie
: APPLICANT: Perna, Nicole T.
: APPLICANT: Plunkett, Guy
: APPLICANT: Welch, Rod
: TITLE OF INVENTION: No. 6365723e1 Sequences of E. coli O157
: NUMBER OF SEQUENCES: 265
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Quarles & Brady
: STREET: 1 South Plinkney Street
: CITY: Madison
: STATE: WI
: COUNTRY: US
: ZIP: 53701-2113

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 168:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-09-453-702B-168

Query Match          21.4%; Score 18; DB 4; Length 80;
Best Local Similarity 56.9%; Pred. No. 5.6e+02;
Matches 33; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY      27  TTTGTGTCCTCTGTTGTGCTAACTTAAAGAGCGGTGTTCTCGAATTGATPAGA 84
          |||||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      75  TTTGTGCTTTCCGCGTGAAGCGGTATATAAATATTGCGAGCTCAATATTAAAGA 18

RESULT 4
US-09-641-638-982/C
; Sequence 982, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bouquelieret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Anick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARCHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 982
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 10-67-340 : polymorphic base C or T
US-09-641-638-982

```



```

: APPLICATION NUMBER: 07/714,131
: FILING DATE: 10-JUNE-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/536,428
: FILING DATE: 11-JUNE-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/964,624
: FILING DATE: 21-OCTOBER-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Barry J. Swanson
: REGISTRATION NUMBER: 33,215
: REFERENCE/DOCKET NUMBER: NEX31/CTP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 793-3333
: TELEFAX: (303) 793-3433
: INFORMATION FOR SEQ ID NO: 317:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 84
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: OTHER INFORMATION: All C's are 2'-F cytosine
: FEATURE:
: OTHER INFORMATION: All U's are 2'-F uracil
: US-08-976-413A-317

Query Match 20.5%; Score 17.2; DB 3; Length 84;
Best Local Similarity 57.4%; Pred. No. 1.1e+03;
Matches 31; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 28 TTGTGTCCTCTGTGTGGCTAACTTTAAGAGCGGTCTTCTCAATTGATA 81
      ||||| | | | | | | | | | | | | | | | | | | | | | |
Db 79 TTGTGAGCCTCTCTGTGCAACAAGCATGAAGTACGCGCACTTTGAGACAACGTGATA 26

RESULT 9
US-09-387-300-33/C
: Sequence 33, Application US/09387300
: Patent No. 6358685
: GENERAL INFORMATION:
: APPLICANT: Wetmur, James G
: APPLICANT: Quattrin, Robin S
: TITLE OF INVENTION: Branch Migration of Nucleotides
: FILE REFERENCE: EN2-49(P)(C)SEQUENCES
: CURRENT APPLICATION NUMBER: US/09/387,300
: CURRENT FILING DATE: 1999-08-31
: EARLIER APPLICATION NUMBER: 08/480,000
: EARLIER FILING DATE: 1995-06-07
: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 33
: LENGTH: 39
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: PMS
: OTHER INFORMATION: oligodeoxynucleotide segment
: US-09-387-300-33

Query Match 20.2%; Score 17; DB 4; Length 39;
Best Local Similarity 69.7%; Pred. No. 1e+03;
Matches 23; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 24 GTGTTTGTCCTCTCTGTGGCTAACTTTAA 56
      ||||| ||||| | | | | | | | | | |
Db 36 GTGTGTGTGTCCCGGGGTGGCCAGATCTCA 4

RESULT 10
US-08-805-918-33
Sequence 33, Application US/08805918

```



```
; Patent No. 5885821
; GENERAL INFORMATION:
; APPLICANT: MAGOTA, Koji
; APPLICANT: MASUDA, Toyofumi
; APPLICANT: SUZUKI, Yuji
; APPLICANT: YABUTA, Masayuki
; TITLE OF INVENTION: PROCESS FOR PRODUCTION OF SECRETORY KEX2
; TITLE OF INVENTION: DERIVATIVES
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22131-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,918
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-073217
; FILING DATE: 04-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-352580
; FILING DATE: 16-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 001560-295
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA"
; US-08-805-918-33

Query Match          20.2%; Score 17; DB 2; Length 42;
Best Local Similarity 63.4%; Pred. No. 1.1e+03;
Matches 26; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 35 TTCCTGTGTGCTTAAGACGGGTCTTCAGAA 75
DB 2 TTCGCGGAGGAGCCTCTGTAAAGGCGTCTTCTGAA 42

RESULT 11
US-08-811-028-21
; Sequence 21, Application US/08811028C
; Patent No. 5891671
; GENERAL INFORMATION:
; APPLICANT: SUZUKI, Yuji
; APPLICANT: MASUDA, Koji
; APPLICANT: MASUDA, Toyofumi
; TITLE OF INVENTION: METHOD FOR CLEAVING CHIMERIC ENZYME USING PROCESSING
; FILE REFERENCE: 001560-294
; CURRENT APPLICATION NUMBER: US/08/811,028C
; CURRENT FILING DATE: 1987-03-04
; EARLIER APPLICATION NUMBER: JP 8-70906
; EARLIER FILING DATE: 1996-03-04
; NUMBER OF SEQ ID NOS: 54
```

```
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 21
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer 507
; US-08-811-028-21

Query Match          20.2%; Score 17; DB 2; Length 42;
Best Local Similarity 63.4%; Pred. No. 1.1e+03;
Matches 26; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 35 TTCCTGTGTGCTTAAGACGGGTCTTCAGAA 75
DB 2 TTCGCGGAGGAGCCTCTGTAAAGGCGTCTTCTGAA 42

RESULT 12
US-08-222-177A-361/C
; Sequence 361, Application US/08222177A
; Patent No. 5582979
; GENERAL INFORMATION:
; APPLICANT: Weber, James L.
; TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
; TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 460
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dewilt Ross & Stevens, S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,177A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/341,562
; FILING DATE: 21-APR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: 09865.601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 831-2100
; TELEFAX: (608) 831-2106
; TELEX:
; INFORMATION FOR SEQ ID NO: 361:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: md114irs
; US-08-222-177A-361

Query Match          20.0%; Score 16.8; DB 1; Length 49;
Best Local Similarity 75.0%; Pred. No. 1.3e+03;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 24 GTGTGTGTGTCTCTGTGTGCTTAAC 51
DB 28 GTGTGTGTGTCTGTCTATGATTAAC 1
```

48 TAACITTAAGAGCGGTGTTCTCAGAA TTGATAAG 83
||||| |
34 TAACTTTAAGAGAGGAGAA TGAATGAATTGGAAG 69

Oy 23 AGTGTGTTGTGTGTTCTGTGTGGCTAATTGAAGCGGTGTTCTCAGA 74
 || | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 8 AGCTTAAGAGTGTACCAACATTGGCAGAAAGTGAGCAGAGTGTTTTACGA 59

```

RESULT 15
US-09-218-444-20
; Sequence 20, Application US/09218444
; Patent No. 6238888
; GENERAL INFORMATION:
; APPLICANT: Gentz, Reiner L.
; APPLICANT: Chopra, Arvind
; APPLICANT: Kaushal, Parveen
; APPLICANT: Spitznagel, Thomas
; APPLICANT: Unsworth, Edward
; APPLICANT: Khan, Fazal
; TITLE OF INVENTION: Keratinocyte Growth Factor-2 Formulations
; FILE REFERENCE: 1488.1030001
; CURRENT APPLICATION NUMBER: US/09/218,444
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: US 60/068,493
; EARLIER FILING DATE: 1997-12-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 76
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-218-444-20

Query Match          20.0%; Score 16.8; DB 4; Length 76;
Best Local Similarity 57.7%; Pred. No. 1.5e+03;
Matches 30; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

Oy 23 AGTGTTCGTGTCCTGCTGCTACTTTAAGAACGGTGTTCACGA 74
|| | | | | | | | | | | | | | | | | |
Db 8 AGCTTAAGAGTGTACCACTTGGCAGAAAGTGAAGAGGTTTACGA 59

Search completed: February 16, 2003, 05:05:03
Job time : 68 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 04:26:52 : Search time 100 Seconds
(without alignments)
427.832 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84
Sequence: 1 actctctctctcagatcag.....gtctctcagatgtataga 84

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 260506

Minimum DB seq length: 0
Maximum DB seq length: 84

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------------------|---------------------|
| C 1 | 20.4 | 24.3 | 80 | 8 US-08-961-888-22 | Sequence 22, Appl |
| C 2 | 20.4 | 24.3 | 81 | 8 US-08-961-888-21 | Sequence 21, Appl |
| C 3 | 18 | 21.4 | 60 | 7 US-08-785-997-23 | Sequence 23, Appl |
| C 4 | 18 | 21.4 | 60 | 9 US-09-387-340-23 | Sequence 23, Appl |
| C 5 | 18 | 21.4 | 80 | 9 US-10-114-170-168 | Sequence 168, Appl |
| C 6 | 17.8 | 21.2 | 53 | 9 US-09-968-561A-339 | Sequence 339, Appl |
| C 7 | 17.8 | 21.2 | 53 | 10 US-09-192-854-201 | Sequence 201, Appl |
| C 8 | 17.4 | 20.7 | 78 | 10 US-09-969-373-1389 | Sequence 1389, Appl |
| C 9 | 17.4 | 20.7 | 74 | 10 US-09-783-590-1818 | Sequence 1818, Appl |
| C 10 | 17.2 | 20.5 | 79 | 10 US-09-864-761-1836 | Sequence 1836, A |
| C 11 | 17.2 | 20.0 | 76 | 10 US-09-853-666-20 | Sequence 20, Appl |
| C 12 | 16.8 | 20.0 | 77 | 10 US-09-864-761-1836 | Sequence 1836, A |
| C 13 | 16.8 | 20.0 | 77 | 10 US-09-969-373-1390 | Sequence 1390, A |
| C 14 | 16.8 | 20.0 | 84 | 10 US-09-864-761-1850 | Sequence 1850, A |
| C 15 | 16.6 | 19.8 | 55 | 10 US-09-943-286-1 | Sequence 1, Appl |
| C 16 | 16.6 | 19.8 | 57 | 10 US-09-923-876-2258 | Sequence 2258, Appl |
| C 17 | 16.4 | 19.5 | 31 | 10 US-09-801-274-424 | Sequence 424, Appl |
| C 18 | 16.4 | 19.5 | 50 | 9 US-10-066-500-30 | Sequence 30, Appl |
| C 19 | 16.2 | 19.3 | 27 | 9 US-09-938-689-20 | Sequence 20, Appl |

| | | | | | |
|------|------|------|----|------------------------|---------------------|
| C 20 | 16.2 | 19.3 | 47 | 9 US-09-853-526-306 | Sequence 306, Appl |
| C 21 | 16.2 | 19.3 | 47 | 10 US-09-901-484A-306 | Sequence 306, Appl |
| C 22 | 16.2 | 19.3 | 60 | 9 US-09-902-941-1987 | Sequence 1987, Appl |
| C 23 | 16.2 | 19.3 | 61 | 9 US-09-999-686-5 | Sequence 5, Appl |
| C 24 | 16.2 | 19.3 | 75 | 10 US-09-864-761-26979 | Sequence 26979, A |
| C 25 | 16.2 | 19.3 | 77 | 10 US-09-864-761-24564 | Sequence 24564, A |
| C 26 | 16.2 | 19.3 | 78 | 10 US-09-732-091-23 | Sequence 23, Appl |
| C 27 | 16.2 | 19.3 | 79 | 9 US-10-015-219-1031 | Sequence 1031, Appl |
| C 28 | 16.2 | 19.3 | 79 | 10 US-09-777-564-1031 | Sequence 1031, Appl |
| C 29 | 16.2 | 19.3 | 83 | 10 US-09-864-761-32212 | Sequence 32212, A |
| C 30 | 16 | 19.0 | 40 | 9 US-09-803-454-43 | Sequence 43, Appl |
| C 31 | 16 | 19.0 | 52 | 9 US-10-056-414-810 | Sequence 810, Appl |
| C 32 | 16 | 19.0 | 61 | 9 US-09-902-941-1986 | Sequence 1986, Appl |
| C 33 | 16 | 19.0 | 63 | 10 US-09-953-321-8 | Sequence 8, Appl |
| C 34 | 16 | 19.0 | 67 | 9 US-09-884-799-32 | Sequence 32, Appl |
| C 35 | 16 | 19.0 | 81 | 10 US-09-983-965-3236 | Sequence 3236, Appl |
| C 36 | 15.8 | 18.8 | 63 | 10 US-09-923-876-686 | Sequence 686, Appl |
| C 37 | 15.8 | 18.8 | 71 | 10 US-09-333-527-8 | Sequence 8, Appl |
| C 38 | 15.8 | 18.8 | 74 | 10 US-09-773-385-4 | Sequence 4, Appl |
| C 39 | 15.8 | 18.8 | 76 | 9 US-10-015-219-197 | Sequence 197, Appl |
| C 40 | 15.8 | 18.8 | 76 | 10 US-09-777-564-197 | Sequence 197, Appl |
| C 41 | 15.8 | 18.8 | 84 | 10 US-09-864-761-21161 | Sequence 21161, A |
| C 42 | 15.6 | 18.6 | 49 | 10 US-09-773-385-3 | Sequence 3, Appl |
| C 43 | 15.6 | 18.6 | 58 | 10 US-09-773-385-2 | Sequence 2, Appl |
| C 44 | 15.4 | 18.3 | 31 | 9 US-09-864-785-2361 | Sequence 2361, Appl |
| C 45 | 15.4 | 18.3 | 45 | 9 US-09-978-295A-307 | Sequence 307, Appl |

ALIGNMENTS

RESULT 1
US-08-961-888-22/C
Sequence 22, Application US/08961888
Patent No. US20010016351A1
GENERAL INFORMATION:
APPLICANT: Padgett, Kerstien
TITLE OF INVENTION: No. US20010016351A1 Vector For Gene Expression
TITLE OF INVENTION: In Prokaryotic And Eukaryotic Systems
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004-2402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: IBM Compatible
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,888
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Auerbach, Jeffrey I
REGISTRATION NUMBER: 32,680
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 383-7451
TELEFAX: 202 383-6610
TELEX:
INFORMATION FOR SEQ. ID NO.: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 base pairs
TYPE: nucleic acid
STRANDEDNESS: single


```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/387,340
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamsen, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: MON-102.0 6018/69242
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)655-1500
; TELEFAX: (312)655-1501
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; POSITION IN GENOME:
; UNITS: bp
;
; US-09-387-340-23

```

```

Query Match      21.4%; Score 18; DB 9; Length 60;
Best Local Similarity 80.8%; Pred. No. 2.6e+03;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy      53 TTAAAGACGGTGTCTCAGAAATG 78
      ||||||| | ||||| ||||| |
Db      43 TTAAAGAGAGCTTTCTTAGAATAG 18

```

```

RESULT 5
US-10-114-170-168/c
; Sequence 168, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blatner, Frederick R.
;            Burland, Valerie
;            Perna, Nicole T.
;            Plunkett, Guy
;            Welch, Rod
; TITLE OF INVENTION: No. US20030023075A1 Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,170
; FILING DATE: 01-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/453,702
; FILING DATE: 03-DEC-1999
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296, 95017
;

```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 168:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 168:
;
; US-10-114-170-168

```

```

Query Match      21.4%; Score 18; DB 9; Length 80;
Best Local Similarity 56.9%; Pred. No. 2.8e+03;
Matches 33; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```

```

Qy      27 TTGTGTGTTCTCGTGTGGCTAATTAAAGACGGTGTCTCAGAAATGATAGA 84
      ||||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db      75 TTGTGTGTTTCGCTGAGCGCTAATATAAATATTGCCAGCTCATATATTAAAGA 18

```

```

RESULT 6
US-09-968-561A-339
; Sequence 339, Application US/09968561A
; Patent No. US20020164642A1
; GENERAL INFORMATION:
; APPLICANT: Tomlinson, Ian M
; APPLICANT: Winter, Gregory
; TITLE OF INVENTION: Method to Screen Phage Display Libraries with Different Ligand
; FILE REFERENCE: 8039/1073B
; CURRENT APPLICATION NUMBER: US/09/968,561A
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 9722131.1
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/065,248
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/066,729
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: PCT/GB98/03135
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 09/511,939
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 339
; LENGTH: 53
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Synthetic PCR primer including degeneracy introduced by the NN
; OTHER INFORMATION: ocons.
;
; US-09-968-561A-339

```

```

Query Match      21.2%; Score 17.8; DB 9; Length 53;
Best Local Similarity 53.3%; Pred. No. 2.9e+03;
Matches 24; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

```

```

Qy      34 GTTCGTGTTGTGCTAATTAAAGACGGTGTCTCAGAAATG 78
      || ||||| | | ||||| | ||||| | ||||| | ||||| |
Db      4 GTTTCGTGCTATACCAABHTAABHBABHAATGCTCTGACTTG 48

```

```

RESULT 7
US-09-192-854-201
; Sequence 201, Application US/09192854
; Patent No. US20020068276A1
; GENERAL INFORMATION:
; APPLICANT: Winter, Greg
; APPLICANT: Tomlinson, Ian
; TITLE OF INVENTION: Methods for Selecting Functional Peptides
; FILE REFERENCE: 3789/72916
;

```



```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (43)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (64)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (68)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (70)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (74)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-1818
```

Query Match 20.5%; Score 17.2; DB 10; Length 74;
Best Local Similarity 59.6%; Pred. No. 5e+03;

Matches 28; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
QY 25 TGTTCGTGTCCTGCTGTGCTTAAGACGGGTTCCTC 71
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 TTTTTCCTTTTTCCTGTCGCTTTTTCACAGCTTTTTCCTC 13
```

RESULT 11

US-09-864-761-18836
; Sequence 18836, Application US/09864761
; Patent No. US20020048763A1

GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; PRIOR FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263,6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 18836

; LENGTH: 79

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC010134.1

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.91

; OTHER INFORMATION: EST_HUMAN HIT: BE966691.2, EVALU6 9.40e-01

; OTHER INFORMATION: NT HIT: AF090444.1, EVALU6 6.30e-01

US-09-864-761-18836

Query Match 20.5%; Score 17.2; DB 10; Length 79;
Best Local Similarity 60.9%; Pred. No. 5.1e+03;

Matches 28; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
QY 11 TCAGGTACAGACAGTGTGTCCTGTCGCTTAAGTAA 56
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 TCAGACACGACACACACTTCTGTCACCTCAGCTTAAGTAA 70
```

RESULT 12

US-09-853-666-20
; Sequence 20, Application US/09853666
; Patent No. US20020016295A1

GENERAL INFORMATION:

; APPLICANT: Gentz, Reiner L.

; APPLICANT: Chopra, Arvind

; APPLICANT: Kauschal, Parveen

; APPLICANT: Spliznagel, Thomas

; APPLICANT: Unsworth, Edward

; APPLICANT: Khan, Fazal

; TITLE OF INVENTION: Keratinocyte Growth Factor-2 Formulations

; FILE REFERENCE: 1488.1030001

; CURRENT APPLICATION NUMBER: US/09/853,666

; CURRENT FILING DATE: 2001-05-14

; PRIOR APPLICATION NUMBER: 09/218,444

; PRIOR FILING DATE: 1998-12-22

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 20

; LENGTH: 76

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-853-666-20

Query Match 20.0%; Score 16.8; DB 10; Length 76;
Best Local Similarity 57.7%; Pred. No. 6.9e+03;

Matches 30; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
QY 23 AGTGTTGTCGTCCTGTCGCTTAAGACGGGTTCACA 74
      || | ||| ||| ||| ||| ||| ||| ||| |||
Db 8 AGCTTAAGAGTGTACACCAATGTCAGAAAGTGACAGAGTTCACGA 59
```

RESULT 13

US-09-969-373-1390/C
; Sequence 1390, Application US/09969373
; Patent No. US2002013852A1

; GENERAL INFORMATION:
; APPLICANT: Efferetz, Roger J.

```

; APPLICANT: Hauge, Brian M.
; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/760,427
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 09/855,768
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 1390
; LENGTH: 77
; TYPE: DNA
; ORGANISM: Glycine max
; US-09-969-373-1390

Query Match          20.0%; Score 16.8; DB 10; Length 77;
Best Local Similarity 57.7%; Pred. No. 6.9e+03;
Matches 30; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 23 AGGTTGCTGTCTCTGCTGCTAAGGCGGTGTTCTCAGA 74
Db 52 AGGTTGCTGTCTCTGCTGCTAAGGCGGTGTTCTCAGA 1

RESULT 14
US-09-864-761-18050
; Sequence 18050, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
```

```

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18050
; LENGTH: 84
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000054.10
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
; OTHER INFORMATION: SWISSPROT HIT: P48903, EVALU 7.10e+00
; US-09-864-761-18050

Query Match          20.0%; Score 16.8; DB 10; Length 84;
Best Local Similarity 75.0%; Pred. No. 7.1e+03;
Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 15 GCACAGCAGCTGTTGCTGCTCTCTGT 42
Db 28 GCACAGCAGCTGTTGCTGCTCTCTGT 55

RESULT 15
US-09-943-286-1
; Sequence 1, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 55
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer T7A(-)4190
; US-09-943-286-1

Query Match          19.8%; Score 16.6; DB 10; Length 55;
Best Local Similarity 71.0%; Pred. No. 7.4e+03;
Matches 22; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 19 AGACAGCTGTTGCTGCTCTCTGTGCGCTA 49
Db 25 AGACAGCTGTTGCTGCTCTCTGTGCGCTA 55
```

Search completed: February 16, 2003, 06:16:22
Job time : 101 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 03:28:49 : Search time 3775 Seconds
(without alignments)
559.462 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84

Sequence: 1 acttcattcttcaggtacag.....gtttcagattgataaga 84

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 13264880

Minimum DB seq length: 0
Maximum DB seq length: 84

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

```

1: /cgn2_6/ptodata/2/pna/US06000.COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06001.COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US06002.COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US06003.COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US06004.COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US06005.COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US06006.COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US06007.COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US06008.COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US06009.COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US06010.COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US06011.COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US06012.COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US06013.COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US06014.COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US06015.COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US06016.COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US06017.COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US06018.COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US06019.COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US06020.COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US06021.COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US06022.COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US06023.COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US06024.COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US06025.COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US06026.COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US06027.COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US06028.COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US06029.COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US06030.COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US06031.COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US06032.COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US06033.COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US06034.COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US06035.COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US06036.COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US06037.COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US06038.COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US06039.COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US06040.COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US06041.COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US06042.COMB.seq:*

```

```

44: /cgn2_6/ptodata/2/pna/US06000.COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US06001.COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US06002.COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US06003.COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US06004.COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US06005.COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US06006.COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US06007.COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US06008.COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US06009.COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US06010.COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US06011.COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US06012.COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US06013.COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US06014.COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US06015.COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US06016.COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US06017.COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US06018.COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US06019.COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US06020.COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US06021.COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US06022.COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US06023.COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US06024.COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US06025.COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US06026.COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US06027.COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US06028.COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US06029.COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US06030.COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US06031.COMB.seq:*
76: /cgn2_6/ptodata/2/pna/US06032.COMB.seq:*
77: /cgn2_6/ptodata/2/pna/US06033.COMB.seq:*
78: /cgn2_6/ptodata/2/pna/US06034.COMB.seq:*
79: /cgn2_6/ptodata/2/pna/US06035.COMB.seq:*
80: /cgn2_6/ptodata/2/pna/US06036.COMB.seq:*
81: /cgn2_6/ptodata/2/pna/US06037.COMB.seq:*
82: /cgn2_6/ptodata/2/pna/US06038.COMB.seq:*
83: /cgn2_6/ptodata/2/pna/US06039.COMB.seq:*
84: /cgn2_6/ptodata/2/pna/US06040.COMB.seq:*
85: /cgn2_6/ptodata/2/pna/US06041.COMB.seq:*

```

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|---------------------|
| 1 | 23.6 | 28.1 | 65 | US-10-032-585-2357 | Sequence 2357, Ap |
| 2 | 23.6 | 28.1 | 65 | US-60-314-050-2357 | Sequence 2357, Ap |
| 3 | 23.6 | 28.1 | 83 | US-60-259-128-1596 | Sequence 1596, Ap |
| 4 | 22 | 26.2 | 47 | PCR-US01-47523-40 | Sequence 40, Appl |
| 5 | 22 | 26.2 | 47 | PCR-US01-47523-42 | Sequence 42, Appl |
| 6 | 21.8 | 26.0 | 80 | US-09-478-821-5377 | Sequence 5377, Ap |
| 7 | 21.8 | 26.0 | 80 | US-09-540-229-147651 | Sequence 147651, Ap |
| 8 | 21.8 | 26.0 | 80 | US-60-115-695-979 | Sequence 979, Ap |
| 9 | 21.2 | 25.2 | 81 | US-09-611-643-5125 | Sequence 5125, Ap |
| 10 | 21.2 | 25.2 | 81 | US-09-654-617-335920 | Sequence 335920, Ap |
| 11 | 21.2 | 25.2 | 81 | US-09-684-016-335920 | Sequence 335920, Ap |
| 12 | 21 | 25.0 | 51 | US-09-755-374A-17078 | Sequence 17078, A |
| 13 | 20.8 | 24.8 | 60 | US-09-908-975-18125 | Sequence 18125, A |
| 14 | 20.8 | 24.8 | 60 | US-09-908-975A-18125 | Sequence 18125, A |
| 15 | 20.8 | 24.8 | 72 | US-60-287-724-18125 | Sequence 18125, A |
| 16 | 20.4 | 24.3 | 47 | PCR-US01-47523-41 | Sequence 41, Appl |
| 17 | 20.4 | 24.3 | 47 | PCR-US01-47523-43 | Sequence 43, Appl |
| 18 | 20.4 | 24.3 | 51 | US-60-213-360-5625 | Sequence 5625, Ap |
| 19 | 20.4 | 24.3 | 66 | US-60-207-458-54398 | Sequence 54398, A |
| 20 | 20.4 | 24.3 | 75 | US-08-957-941-210 | Sequence 210, App |
| 21 | 20.4 | 24.3 | 75 | US-60-029-494-210 | Sequence 210, App |

```

c 22 20.4 24.3 80 13 US-08-961-888-22 Sequence 22, Appl
c 23 20.4 24.3 81 13 US-08-961-888-21 Sequence 21, Appl
c 24 20.2 24.0 50 11 US-08-781-986A-5149 Sequence 5149, Ap
c 25 20.2 24.0 50 13 US-08-956-171E-5149 Sequence 5149, Ap
c 26 20.2 24.0 50 13 US-08-956-171B-5149 Sequence 5149, Ap
c 27 20.2 24.0 50 13 US-08-956-171C-5149 Sequence 5149, Ap
c 28 20.2 24.0 51 13 US-08-956-171E-5149 Sequence 5149, Ap
c 29 20.2 24.0 51 29 US-09-755-374A-27189 Sequence 27189, A
c 30 20.2 24.0 51 29 US-09-755-374A-27190 Sequence 27190, A
c 31 20.2 24.0 51 65 US-60-213-359-752 Sequence 752, App
c 32 20.2 24.0 51 71 US-60-324-185-12891 Sequence 12891, A
c 33 20.2 24.0 60 24 US-09-628-859-1305 Sequence 1305, Ap
c 34 20.2 24.0 66 17 US-09-306-349-30539 Sequence 30539, A
c 35 20.2 24.0 66 36 US-09-960-481-30539 Sequence 30539, A
c 36 20.2 24.0 77 12 US-08-826-438-2395 Sequence 2395, Ap
c 37 20.2 24.0 77 13 US-08-951-197-2395 Sequence 2395, Ap
c 38 20.2 24.0 77 13 US-08-826-438-2395 Sequence 2395, Ap
c 39 20.2 24.0 77 20 US-09-539-806-34124 Sequence 34124, A
c 40 20 23.8 50 29 US-09-755-374A-11319 Sequence 11319, A
c 41 20 23.8 51 29 US-09-755-374A-19257 Sequence 19257, A
c 42 20 23.8 78 14 US-09-070-695A-2493 Sequence 2493, Ap
c 43 20 23.8 81 15 US-09-540-229-23150 Sequence 23150, A
c 44 20 23.8 81 15 US-09-103-841-1102 Sequence 1102, Ap
c 45 20 23.8 81 20 US-09-534-853-4977 Sequence 4977, Ap
```

ALIGNMENTS

```

RESULT 1
; Sequence 2357, Application US/10032585
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2357
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-032-585-2357

Query Match          28.1%; Score 23.6; DB 38; Length 65;
Best Local Similarity 64.8%; Pred. No. 1.4e+03;
Matches 35; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Oy 14 GGTACAGACAGTGTGTGCTGCTGTTGGCTAACTTTAAGAGCGGTGTT 67
      ||||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| |||
Db 5 GGTAAAGACATTGTGTGTGATTGATTCGATAGTATATTAATGAAGAGAT 58

RESULT 2
; Sequence 2357, Application US/60314050
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; APPLICANT: Ohlsen, Kari L.
; TITLE OF INVENTION: GENE DISRUPTION METHODOLOGIES FOR DRUG TARGET DISCOVERY
; FILE REFERENCE: 10182-013-888
; CURRENT APPLICATION NUMBER: US/60/314,050
; CURRENT FILING DATE: 2001-08-22
; NUMBER OF SEQ ID NOS: 7834
; SOFTWARE: PatentIn version 3.1
```

```

; SEQ ID NO 2357
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Candida albicans
US-60-314-050-2357
```

```

Query Match          28.1%; Score 23.6; DB 75; Length 65;
Best Local Similarity 64.8%; Pred. No. 1.4e+03;
Matches 35; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```

Oy 14 GGTACAGACAGTGTGTGCTGCTGTTGGCTAACTTTAAGAGCGGTGTT 67
      ||||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| |||
Db 5 GGTAAAGACATTGTGTGTGATTGATTCGATAGTATATTAATGAAGAGAT 58
```

```

RESULT 3
; Sequence 1596, Application US/60259128
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: GRACE1
; CURRENT APPLICATION NUMBER: US/60/259,128
; CURRENT FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 5129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1596
; LENGTH: 83
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA primer
US-60-259-128-1596
```

```

Query Match          28.1%; Score 23.6; DB 69; Length 83;
Best Local Similarity 64.8%; Pred. No. 1.5e+03;
Matches 35; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```

Oy 14 GGTACAGACAGTGTGTGCTGCTGTTGGCTAACTTTAAGAGCGGTGTT 67
      ||||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| |||
Db 5 GGTAAAGACATTGTGTGTGATTGATTCGATAGTATATTAATGAAGAGAT 58
```

```

RESULT 4
; Sequence 40, Application PC/TUS0147523
; GENERAL INFORMATION:
; APPLICANT: Cold Spring Harbor Laboratory
; TITLE OF INVENTION: Chimeric Molecules to Modulate Gene Expression
; FILE REFERENCE: 1020-3 PCT
; CURRENT APPLICATION NUMBER: PCT/US01/47523
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: U.S. 60/304,182
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: primer
; NAME/KEY: misc-feature
; LOCATION: (1)..(47)
; OTHER INFORMATION: primer
PCT-US01-47523-40
```

```

Query Match          26.2%; Score 22; DB 1; Length 47;
```



```

: APPLICANT: MASSEMAN, ALON
: APPLICANT: MINTZ, EJI
: APPLICANT: MINTZ, LIAT
: APPLICANT: FAIGLER, SIMCHON
: TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICED
: TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
: FILE REFERENCE: 36688-0005
: CURRENT APPLICATION NUMBER: US/09/908,975
: CURRENT FILING DATE: 2001-07-20
: PRIOR APPLICATION NUMBER: US 60/287,724
: PRIOR FILING DATE: 2001-05-02
: PRIOR APPLICATION NUMBER: US 60/221,607
: PRIOR FILING DATE: 2000-07-28
: NUMBER OF SEQ ID NOS: 3237
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 18125
: LENGTH: 60
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-908-975-18125

```

| | | | | | | | |
|-----------------------|--------|--------------|----------|------------|-----|--------|-----|
| Query Match | 24.88; | Score | 20.8; | DB | 34; | Length | 60; |
| Best Local Similarity | 64.68; | Pred. No. | 1.2e+04; | | | | |
| Matches | 31; | Conservative | 0; | Mismatches | 17; | Indels | 0; |
| | | | | | | Gaps | 0; |

OY 14 GGTACAGACAGTGTTCGTGTCCTGTTCGGCTTAACCTTAAGAACC 61
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 GGTTTGGAAGGGAATGTTGGTCCTTTTGTCACACAATTAAAGATGC 60

RESULT 14

```

US-09-908-975A-18125
; Sequence 18125, Application US/09908975A
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchaon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICED
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0006
; CURRENT APPLICATION NUMBER: US/09/908, 975A
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18125
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975A-18125

```

| | | | | |
|--------------------------|--------|--------------------|-----------|------------|
| Query Match | 24.88; | Score 20.8; | DB 34; | Length 60; |
| Best Local Similarity | 64.68; | Pred. No. 1.2e+04; | | |
| Matches 31; Conservative | 0; | Mismatches 17; | Indels 0; | Gaps 0; |

QY 14 GGTACAGACAGTGTGTGTCCTGTTGGCCTTAACCTTAAGAAGC 61
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 GGT TTGGAAGGGAATGTTGGTGCCCTTTGTGCCACAAGTTAAGATGC 60

RESULT 15

US-60-287-724-18125
: Sequence 18125, Application US/60287724
: GENERAL INFORMATION:
: APPLICANT: SHOSMAN, AVI
: APPLICANT: WASSERMAN, ALON
: APPLICANT: MINTZ, ELI
: APPLICANT: MINTZ, LIAT

```

APPLICANT: FAIGLER, Simchon
:
: TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPL
:
: TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
:
: FILE REFERENCE: 36688-0004
:
: CURRENT APPLICATION NUMBER: US/60/287,724
:
: CURRENT FILING DATE: 2001-05-02
:
: NUMBER OF SEQ ID NOS: 32337
:
: SOFTWARE: Patentin version 3.0
:
: SEQ ID NO 18125
:
: LENGTH: 60
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
US-60-287-724--18125

```

| | | | | |
|--------------------------|--------|--------------------|-----------|------------|
| Query Match | 24.88; | Score 20.8; | DB 72; | Length 60; |
| Best Local Similarity | 64.68; | Pred. No. 1.2e+04; | | |
| Matches 31; Conservative | 0; | Mismatches 17; | Indels 0; | Gaps 0; |

OY 14 GGTACAGACAGCTGTTTGTGTGTCCCTGTTGTGGCTAACTTTAAGAACC 61
 ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 GGTTTGGAAGGAAATGTTGGTGCCCTTTTGTGCCACAAGTTAAGATGC 60

Search completed: February 16, 2003, 06:08:13
Job time : 3777 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 03:33:37 : Search time 363 Seconds

(without alignments)
344,867 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84
Sequence: 1 actctatcttcagtgacag.....gtttccagattgataaga 84

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2773584 seqs, 745158349 residues

Total number of hits satisfying chosen parameters: 4588940

Minimum DB seq length: 0
Maximum DB seq length: 84

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--|
| C 1 | 20.2 | 24.0 | 50 | 6 | US-10-329-624-5149 Sequence 5149, App |
| C 2 | 19.2 | 22.9 | 60 | 1 | PCT-US02-38147-180 Sequence 180, App |
| C 3 | 18.6 | 22.1 | 77 | 6 | US-10-293-338-4769 Sequence 4769, App |
| C 4 | 18.4 | 21.9 | 50 | 6 | US-10-325-899-7355 Sequence 7355, App |
| C 5 | 18.4 | 21.9 | 61 | 6 | US-10-349-508-66 Sequence 66, Appl |
| C 6 | 18.2 | 21.4 | 66 | 6 | US-10-310-188-66563 Sequence 66563, A |
| C 7 | 18 | 21.4 | 81 | 5 | US-09-331-113-24166 Sequence 24166, A |
| C 8 | 17.8 | 21.2 | 55 | 5 | US-09-968-744A-339 Sequence 339, App |
| C 9 | 17.6 | 21.0 | 25 | 7 | US-60-427-808-910940 Sequence 910940, |
| C 10 | 17.6 | 21.0 | 25 | 7 | US-60-427-836-654815 Sequence 654815, |
| C 11 | 17.6 | 21.0 | 50 | 6 | US-10-325-899-1526 Sequence 1526, App |
| C 12 | 17.6 | 21.0 | 51 | 5 | US-09-513-999C-30195 Sequence 30195, A |
| C 13 | 17.6 | 21.0 | 82 | 5 | PCT-US02-38147-178 Sequence 178, App |
| C 14 | 17.6 | 21.0 | 82 | 5 | US-09-531-113-25646 Sequence 25646, A |
| C 15 | 17.4 | 20.7 | 65 | 6 | US-10-294-040-865 Sequence 865, App |
| C 16 | 17.4 | 20.7 | 81 | 6 | US-10-310-156-508 Sequence 508, App |
| C 17 | 17.4 | 20.7 | 84 | 6 | US-09-331-113-18049 Sequence 18049, A |
| C 18 | 17.2 | 20.5 | 64 | 6 | US-10-293-338-3246 Sequence 3246, App |
| C 19 | 17.2 | 20.5 | 80 | 6 | US-10-329-624-4896 Sequence 4896, App |
| C 20 | 17 | 20.2 | 25 | 7 | US-60-427-808-721196 Sequence 721196, |
| C 21 | 17 | 20.2 | 25 | 7 | US-60-427-808-721196 Sequence 721196, |
| C 22 | 17 | 20.2 | 25 | 7 | US-60-427-808-739340 Sequence 739340, |
| C 23 | 17 | 20.2 | 50 | 6 | US-10-325-899-1164 Sequence 1164, App |
| C 24 | 17 | 20.2 | 64 | 6 | US-10-303-778-4485 Sequence 4485, App |
| C 25 | 17 | 20.2 | 64 | 6 | US-10-310-188-9507 Sequence 9507, App |
| C 26 | 16.8 | 20.0 | 25 | 7 | US-60-417-190-70128 Sequence 70128, A |

| | | | | | |
|------|------|------|----|---|---|
| 27 | 16.8 | 20.0 | 25 | 7 | US-60-417-190-70129 Sequence 70129, A |
| C 28 | 16.8 | 20.0 | 25 | 7 | US-60-427-808-42666 Sequence 42666, A |
| C 29 | 16.8 | 20.0 | 25 | 7 | US-60-427-836-223437 Sequence 223437, App |
| C 30 | 16.8 | 20.0 | 35 | 6 | US-10-142-283-117 Sequence 117, App |
| C 31 | 16.8 | 20.0 | 35 | 6 | US-10-325-899-3362 Sequence 3362, App |
| C 32 | 16.8 | 20.0 | 51 | 6 | US-10-316-957-1630 Sequence 1630, App |
| C 33 | 16.8 | 20.0 | 54 | 5 | US-09-852-903C-30 Sequence 30, Appl |
| C 34 | 16.8 | 20.0 | 56 | 5 | US-09-852-903C-31 Sequence 31, Appl |
| C 35 | 16.8 | 20.0 | 56 | 6 | US-10-329-624-2540 Sequence 2540, App |
| C 36 | 16.8 | 20.0 | 58 | 5 | US-09-852-903C-32 Sequence 32, Appl |
| C 37 | 16.8 | 20.0 | 60 | 5 | US-09-852-903C-33 Sequence 33, Appl |
| C 38 | 16.8 | 20.0 | 62 | 5 | US-09-852-903C-34 Sequence 34, Appl |
| C 39 | 16.8 | 20.0 | 64 | 5 | US-09-852-903C-35 Sequence 35, Appl |
| C 40 | 16.8 | 20.0 | 65 | 7 | US-60-434-832-2107 Sequence 2107, App |
| C 41 | 16.8 | 20.0 | 66 | 5 | US-09-852-903C-36 Sequence 36, Appl |
| C 42 | 16.8 | 20.0 | 68 | 5 | US-09-852-903C-37 Sequence 37, Appl |
| C 43 | 16.8 | 20.0 | 70 | 5 | US-09-852-903C-38 Sequence 38, Appl |
| C 44 | 16.8 | 20.0 | 72 | 6 | US-10-316-957-1390 Sequence 1390, App |
| C 45 | 16.8 | 20.0 | 84 | 6 | US-10-203-138A-6444 Sequence 6444, App |

ALIGNMENTS

RESULT 1
US-10-329-624-5149/C
Sequence 5149, Application US/10329624
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/329, 624
FILING DATE: 27-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5149:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 5149:
US-10-329-624-5149

Query Match 24.0%; Score 20.2; DB 6; Length 50;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 4 TCATTCTTCAGGACAGACAGGTGTTGTCCTGCTGTGTG 45
DB 46 TCTTCTCCGGTACTTAAAGATGTTTCAGTTCCTCCGGGTGTG 5

RESULT 2

PCT-US02-38147-180
; Sequence 180, Application PC/TUS0238147
; GENERAL INFORMATION:
; APPLICANT: Cell-Matrix, Inc.
; APPLICANT: Watkins, Jeffrey D.
; APPLICANT: Huse, William D.
; APPLICANT: Tang, Ying
; APPLICANT: Brooks, Daniel
; APPLICANT: Brooks, Peter C.
; TITLE OF INVENTION: Humanized Collagen Antibodies and
; FILE REFERENCE: PP-IX 5532
; CURRENT APPLICATION NUMBER: PCT/US02/38147
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: US 09/995,529
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 180
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
PCT-US02-38147-180

Query Match 22.9%; Score 19.2; DB 1; Length 60;
Best Local Similarity 61.4%; Pred. No. 3e+03;
Matches 27; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

OY 38 CTTGTGCGTACTTAAGAACGGGTGTTCTCAGCAATTGATA 81
DB 1 CTTGTGCGTACTTAAGAACGGGTGTTCTCAGCAATTGATA 44

RESULT 3

US-10-293-338-4769/c
; Sequence 4769, Application US/10293338
; GENERAL INFORMATION:
; APPLICANT: Rosetagenomics LTD
; TITLE OF INVENTION: BIOINFORMATIALLY DETECTABLE GROUP OF NOVEL REGULATORY GENES AND
; FILE REFERENCE: 45282
; CURRENT APPLICATION NUMBER: US/10/293,338
; CURRENT FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 8785
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4769
; LENGTH: 77
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-293-338-4769

Query Match 22.1%; Score 18.6; DB 6; Length 77;
Best Local Similarity 57.9%; Pred. No. 5e+03;
Matches 33; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

OY 9 CTTGAGGTACAGACAGGTGTTGCTGCTGTTGAGGAGGAGG 65
DB 75 CTGACGACAGAGAGGCTTTGTGCAACACTTATTGGAGAAAGATTACACAGCGGTG 19

RESULT 4
US-10-325-899-7355
; Sequence 7355, Application US/10325899
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Ly, Ngoc
; APPLICANT: Woodward, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING TRANSPL

; FILE REFERENCE: 506612000122
; CURRENT APPLICATION NUMBER: US/10/325,899
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 9966
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7355
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-325-899-7355

Query Match 21.9%; Score 18.4; DB 6; Length 50;
Best Local Similarity 78.6%; Pred. No. 5.5e+03;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 55 AAGAACGGGTGTTCTCAGCAATTGATA 82
DB 3 AGGAAGCGGCTTCTCTCAATTGATA 30

RESULT 5

US-10-349-508-66/c
; Sequence 66, Application US/10349508
; GENERAL INFORMATION:
; APPLICANT: Lassner, Michael
; APPLICANT: Post-Beltemmiller, Dusty
; APPLICANT: Savidge, Beth
; APPLICANT: Weiss, James
; TITLE OF INVENTION: Nucleic Acid Sequences Involved in
; FILE REFERENCE: 17133/02/US
; CURRENT APPLICATION NUMBER: US/10/349,508
; CURRENT FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: 60/129,899
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/146,461
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 61
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: oligonucleotide
US-10-349-508-66

Query Match 21.9%; Score 18.4; DB 6; Length 61;
Best Local Similarity 69.4%; Pred. No. 5.6e+03;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 41 TTGTGCTACTTAAGAACGGGTGTTCTCAGCAAT 76
DB 53 TTGTGCTACTTAAGAACGGGTGTTCTCAGCAAT 18

RESULT 6

```

      PRIOR APPLICATION NUMBER: US 09/511,939
      PRIOR FILING DATE: 2000-02-24
      NUMBER OF SEQ ID NOS: 350
      SOFTWARE: PatentIn version 3.1
      SEQ ID NO 339
      LENGTH: 53
      TYPE: DNA
      ORGANISM: Homo sapiens
      FEATURE:
      NAME/KEY: misc_feature
      OTHER INFORMATION: Synthetic PCR primer including degeneracy introduced by the NNN
      US-09-968-744A-339

Query Match                                21.2%; Score 17.8; DB 5; Length 53;
Best Local Similarity 53.3%; Pred.No.8.7e+03;
Matches 24; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY    34 GTTCTGTGTCGCTAACTTTAGACGGCGTGTCTCAGATTG 78
      ||| ||| | | | : | | : | | : | | : | | | | | |
Db     4 GTTCTGCTGATACCABHTAAABHBABHBAHTGCTTGACTTG 48

RESULT 9
US-60-427-808-910940
; Sequence 910940; Application US/60427808
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528
; CURRENT APPLICATION NUMBER: US/60/427,808
; CURRENT FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 910940
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; US-60-427-808-910940

Query Match                                21.0%; Score 17.6; DB 7; Length 25;
Best Local Similarity 83.3%; Pred.No.9.1e+03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY    24 GTGTTGTGTCGTCCTGTTGTGC 47
      ||||| |||| | | | | | | | |
Db     2 GTGTTCTGTGTCCTCGTGTGAC 25

RESULT 10
US-60-427-836-654815
; Sequence 654815; Application US/60427836
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Rat
; FILE REFERENCE: 3527
; CURRENT APPLICATION NUMBER: US/60/427,836
; CURRENT FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 699466
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 654815
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-60-427-836-654815

Query Match                                21.0%; Score 17.6; DB 7; Length 25;
Best Local Similarity 83.3%; Pred.No.9.1e+03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY    25 TGTTCGTGTCGTCCTGTTGTGCCT 48
      ||||| | | | | | | | | | |
Db     1 TGTTCATTTTTCCTGTTGTGCCT 24

```



```
; NUMBER OF SEQ ID NOS: 1102
; SOFTWARE: Proprietary
; SEQ ID NO 865
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Helicobacter pylori 26695 complete genome.
; FEATURE:
; LOCATION: (1288490)...(1288554)
; OTHER INFORMATION: Chromosome = 1 Strand = negative ConnectonObjectNumber = 1302
US-10-294-040-865

Query Match      20.7%; Score 17.4; DB 6; Length 65;
Best Local Similarity 68.6%; Pred. No. 1.2e+04;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 49 AACCTTAAGACGCGGTGTTCTCAGAAATGATTAAG 83
   ||||||| ||||| ||| || |||
Db 18 AATTTTAAAAAGCGTTTAAAAAGCGTTTAAAG 52
```

Search completed: February 16, 2003, 06:14:29
Job time : 364 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 03:26:03 ; Search time 2204 Seconds

(without alignments)
617,251 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84
Sequence: 1 actcattcttcaggtacag.....gtttctcagaattgataaga 84

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 262628

Minimum DB seq length: 0
Maximum DB seq length: 84

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST :
1: em_estda:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 23 | 27.4 | 78 | 17 | AZ407089 1M0176C10 |
| 2 | 22.4 | 26.7 | 73 | 17 | BH864002 |
| 3 | 21.2 | 25.2 | 77 | 9 | A1566444 |
| 4 | 20.4 | 24.3 | 80 | 10 | AV542665 |
| 5 | 20.4 | 24.3 | 81 | 17 | AZ799436 |
| 6 | 20.2 | 24.0 | 77 | 9 | AA097767 |

| Result | Score | Query Match | Length | DB ID | Description |
|--------|-------|-------------|--------|-------|-------------|
| 7 | 19.6 | 23.3 | 77 | 10 | AM179752 |
| 8 | 19.6 | 23.3 | 79 | 17 | AL766138 |
| 9 | 19.6 | 23.3 | 80 | 9 | AA167266 |
| 10 | 19.4 | 23.1 | 73 | 17 | CNS01V43 |
| 11 | 19.2 | 22.9 | 35 | 17 | BH862300 |
| 12 | 19.2 | 22.9 | 35 | 17 | BH862311 |
| 13 | 19.2 | 22.9 | 71 | 17 | AZ388661 |
| 14 | 19.2 | 22.9 | 76 | 9 | A1051826 |
| 15 | 19.2 | 22.9 | 78 | 12 | BG409184 |
| 16 | 19.2 | 22.6 | 61 | 17 | B05482 |
| 17 | 19.2 | 22.6 | 83 | 9 | AA584879 |
| 18 | 18.8 | 22.4 | 66 | 17 | CNS0012P |
| 19 | 18.8 | 22.4 | 71 | 17 | L48783 |
| 20 | 18.8 | 22.4 | 82 | 9 | AA684425 |
| 21 | 18.6 | 22.1 | 49 | 17 | AZ586447 |
| 22 | 18.6 | 22.1 | 61 | 9 | A1735552 |
| 23 | 18.6 | 22.1 | 70 | 9 | AA940574 |
| 24 | 18.6 | 22.1 | 76 | 9 | AA509646 |
| 25 | 18.6 | 22.1 | 76 | 17 | AZ321896 |
| 26 | 18.6 | 22.1 | 78 | 14 | B0753820 |
| 27 | 18.6 | 22.1 | 82 | 9 | A1798019 |
| 28 | 18.4 | 21.9 | 46 | 9 | AA522160 |
| 29 | 18.4 | 21.9 | 58 | 9 | A1006125 |
| 30 | 18.4 | 21.9 | 69 | 9 | A1157857 |
| 31 | 18.4 | 21.9 | 73 | 17 | BH848460 |
| 32 | 18.4 | 21.9 | 73 | 17 | BH848461 |
| 33 | 18.4 | 21.9 | 80 | 14 | T82935 |
| 34 | 18.4 | 21.9 | 83 | 13 | BM092457 |
| 35 | 18.2 | 21.7 | 81 | 10 | AV963847 |
| 36 | 18.2 | 21.7 | 82 | 9 | AU173692 |
| 37 | 18.2 | 21.7 | 82 | 17 | AZ653201 |
| 38 | 18.2 | 21.4 | 43 | 17 | AZ660216 |
| 39 | 18.2 | 21.4 | 52 | 10 | BE320467 |
| 40 | 18.2 | 21.4 | 75 | 12 | BG099465 |
| 41 | 18.2 | 21.4 | 75 | 17 | AZ623204 |
| 42 | 18.2 | 21.4 | 76 | 9 | AA681911 |
| 43 | 18.2 | 21.4 | 78 | 9 | A1492413 |
| 44 | 18.2 | 21.4 | 81 | 14 | B0666314 |
| 45 | 18.2 | 21.4 | 81 | 17 | B37815 |

ALIGNMENTS

RESULT 1
AZ407089
LOCUS
DEFINITION
1M0176C10R Mouse 10kb plasmid UGCG1M library Mus musculus genomic clone UGCG1M0176C10 R, DNA sequence.
ACCESSION
AZ407089
VERSION
AZ407089.1
KEYWORDS
GT:10531102
SOURCE
GSS.
ORGANISM
house mouse.
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathia; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 78)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Relliy,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weis,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

| REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|-----------|---|---|--------------------|---------|
| 1 | (pages 1 to 77) | Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sclurognathia: Muridae: Murinae: Mus. | | |
| | Marin, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R. | The WashU-HMI Mouse EST Project | Unpublished (1996) | |
| | contact: Marria M/Mouse EST Project | | | |
| | WashU-HMI Mouse EST Project | | | |
| | Washington University School of Medicine | | | |

| | | | | | | | |
|-----------------------|-------|------------|------------|--------|---|--------|----|
| Query Match | 23.3% | Score | 19.6 | DB | 9 | Length | 80 |
| Best Local Similarity | 57.1% | Pred. | No.3.5e+04 | | | | |
| Matches | 28 | Mismatches | 21 | Indels | 0 | Gaps | 0 |

| SOURCE ORGANISM | thale cress, <i>Arabidopsis thaliana</i> |
|-----------------|---|
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 4 |
| 5 | 5 |
| 6 | 6 |
| 7 | 7 |
| 8 | 8 |
| 9 | 9 |
| 10 | 10 |
| 11 | 11 |
| 12 | 12 |
| 13 | 13 |
| 14 | 14 |
| 15 | 15 |
| 16 | 16 |
| 17 | 17 |
| 18 | 18 |
| 19 | 19 |
| 20 | 20 |
| 21 | 21 |
| 22 | 22 |
| 23 | 23 |
| 24 | 24 |
| 25 | 25 |
| 26 | 26 |
| 27 | 27 |
| 28 | 28 |
| 29 | 29 |
| 30 | 30 |
| 31 | 31 |
| 32 | 32 |
| 33 | 33 |
| 34 | 34 |
| 35 | 35 |
| 36 | 36 |
| 37 | 37 |
| 38 | 38 |
| 39 | 39 |
| 40 | 40 |
| 41 | 41 |
| 42 | 42 |
| 43 | 43 |
| 44 | 44 |
| 45 | 45 |
| 46 | 46 |
| 47 | 47 |
| 48 | 48 |
| 49 | 49 |
| 50 | 50 |
| 51 | 51 |
| 52 | 52 |
| 53 | 53 |
| 54 | 54 |
| 55 | 55 |
| 56 | 56 |
| 57 | 57 |
| 58 | 58 |
| 59 | 59 |
| 60 | 60 |
| 61 | 61 |
| 62 | 62 |
| 63 | 63 |
| 64 | 64 |
| 65 | 65 |
| 66 | 66 |
| 67 | 67 |
| 68 | 68 |
| 69 | 69 |
| 70 | 70 |
| 71 | 71 |
| 72 | 72 |
| 73 | 73 |
| 74 | 74 |
| 75 | 75 |
| 76 | 76 |
| 77 | 77 |
| 78 | 78 |
| 79 | 79 |
| 80 | 80 |
| 81 | 81 |
| 82 | 82 |
| 83 | 83 |
| 84 | 84 |
| 85 | 85 |
| 86 | 86 |
| 87 | 87 |
| 88 | 88 |
| 89 | 89 |
| 90 | 90 |
| 91 | 91 |
| 92 | 92 |
| 93 | 93 |
| 94 | 94 |
| 95 | 95 |
| 96 | 96 |
| 97 | 97 |
| 98 | 98 |
| 99 | 99 |
| 100 | 100 |

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 35)

TITLE

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)

JOURNAL
COMMENT

Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

FEATURES
source

This is single pass sequence recovered from the left border of TDNA.
Class: TDNA tagged.
Location/Qualifiers

1..35
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SAIK_089314"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT
ORIGIN

11 a 6 c 4 g 14 t

Query Match 22.9%; Score 19.2; DB 17; Length 35;
Best Local Similarity 75.0%; Pred. No. 4.2e+04;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
DB 1 TAACATACACAGCGCTTTTCTCATATTGA 32

RESULT 12
LOCUS

BH862311 35 bp DNA linear GSS 05-AUG-2002
SAIK_089333 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SAIK_089333, DNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BH862311.1 GI:22097637
GSS.
thale cress.
Arabidopsis thaliana

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 35)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)

TITLE

Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

JOURNAL
COMMENT

This is single pass sequence recovered from the left border of

FEATURES
source

TDNA.
Class: TDNA tagged.
Location/Qualifiers

1..35
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SAIK_089333"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT
ORIGIN

11 a 6 c 4 g 14 t

Query Match 22.9%; Score 19.2; DB 17; Length 35;
Best Local Similarity 75.0%; Pred. No. 4.2e+04;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
DB 1 TAACATACACAGCGCTTTTCTCATATTGA 32

RESULT 13
LOCUS

AZ388861 71 bp DNA linear GSS 02-OCT-2000
1M0149002F Mouse 10kb plasmid U0GC1M library Mus musculus genomic clone U0GC1M0149002 F, DNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AZ388861.1 GI:10502569
GSS.
house mouse.
Mus musculus

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus. 1 (bases 1 to 71)
Dunn,D., Aoyagi,A., Barber,M., Beecorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)

JOURNAL
COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0149 row: O column: 02
Seq primer: CGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 71.
Location/Qualifiers

1..71
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U0GC1M0149002"
/clone_lib="Mouse 10kb plasmid U0GC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PMD420v; Purified genomic DNA from M. musculus C57BL/6J (male); was obtained from the Jackson Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA

vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. RNA from sporadic parathyroid adenomas was kindly provided by Dr. Stephen Marx, National Institute of Diabetes and Digestive and Kidney Diseases, NIH."

a 15 c 6 g 20 t

22.9%; Score 19.2; DB 9; Length 76;
Ally 56.2%; Pred. No. 4.7e+04;
Conservative 0; Mismatches 28; Indels 0; Gaps 0;

GTGACGACACAGTGTGTTGTCTTCTGTTGCGCTTAACGACGGCT 64
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 74 CACTGTTTGGGAGAGTCTTTTCTTCTTCTT

| | | | |
|----|----|------|----|
| 24 | 3 | 1111 | 33 |
| Db | 14 | GTTT | 11 |

| | | | |
|-----------|-------|------|-----------------|
| RESULT 15 | | | |
| BG409184 | | | |
| LOCUS | 78 bp | mRNA | linear |
| BG409184 | | | EST 13-MAR-2001 |

| | | |
|-----------|------------------------|-------------|
| ACCESSION | BG409184 | GI:13315529 |
| VERSION | BG409184.1 | |
| KEYWORDS | EST. | |
| SOURCE | Physcomitrella patens. | |
| ORGANISM | Physcomitrella patens | |

REFERENCE
AUTHORS
1 (bases 1 to 78)
Quattrano, R., Bashirades, S., Cove, D., Cumming, A., Knight, C., Clifton
S., Marra, M., Hillier, L., Pape, D., Martin, J., Wolfe, T., Underwood

K., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T.,
Stepcoe, M., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.

TITLE , Waterston, R. and Wilson, R.
Leeds/Wash U Moss EST Project

JOURNAL COMMENT Unpublished (1999)
Contact: Ralph Quatrano
Ileeds/Wash U Moss Est Project

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
Libraries were constructed by Dr. Stavros Bashlaredes as part of the phycocomitrella EST program (PEP) at the Univ. of Leeds (UK) and Washington Univ. in St. Louis (USA) DNA sequencing by: Washington

University Genome Sequencing Center For information on obtaining a clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
putative full length read
vector to vector length is 79

Seq primer: -40RP from Gibco.
Location/Qualifiers

| | |
|--------|--------|
| source | 1. .78 |
|--------|--------|

```
/organism="Phycomitrella patens"  
/db_xref="taxon:3218"
```

```
/clone="PEP_SOURCE_ID:PPG_CopyA-120912"  
/clone_lib="Moss EST library PPG"
```

/tissue_type="gametophore: 30 day old tissue, ammonium-grown"

```
/lab_host="DH10B"  
/note="Vector: Construction of the cDNA library was
```

performed by Dr. W. Gregg Clark using a modification of the cDNA synthesis protocol developed in the laboratory of

Dr. Michael Lovett by Dr. Yulia Korshunova (personal communication). First polyA + RNA was isolated from total

gametophore RNA using oligo dT magnetic beads. Following this, first strand cDNA synthesis was performed on the

bead-bound polyA + RNAs, during which an oligonucleotide anchor sequence was incorporated onto the 5'-ends of the cDNA. PCR amplification was then used to synthesize the second strand, to amplify the double stranded DNA, and to incorporate dUTP containing sequences into the ends of the double stranded cDNA. This DNA was size selected and cloned into pAMP1 using the CloneAMP pAMP1 System (Life Technologies, GibcoBRL) for cloning amplification products by a non-restriction site dependent process. The cloning was directional based on sequence asymmetry introduced at the ends during PCR amplification. The 3' cDNA ends are proximal to the NotI site of the multiple cloning site in pAMP1. This annealing mixture was transformed into chemically competent DH10B cells and selected for ampicillin resistant growth. The resulting clones (about 330,000) were pooled to make the library."

| | | | | |
|------------|------|------|-----|------|
| BASE COUNT | 27 a | 15 c | 9 g | 27 t |
| ORIGIN | | | | |

ORIGIN

| | | | | |
|--------------------------|--------|--------------------|-----------|------------|
| Query Match | 22.9%; | Score 19.2; | DB 12; | Length 78; |
| Best Local Similarity | 58.9%; | Pred. NO. 4.7e+04; | | |
| Matches 33; Conservative | 0; | Mismatches 23; | Indels 0; | Gaps 0; |

Qy 29 TGTGTTCCCTGTTGTGGCAACTTAAAGACGGTGTTTCAGAAATGATAAGA 84
| | | | | | | | | |
Db 18 TCTGTGTTCCATTGAGCTTAACCTATTCCACCCTGCCTTGTGTACTGTTAAAAA 73

Search completed: February 16, 2003, 05:03:46
Job time : 2208 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 11:58:34 ; Search time 2188.78 Seconds
(without alignments)
1361.478 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 184

Sequence: 1 attcatctctcagtagcag.....ccttagtagtgagcattga 184

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: qb_estl:*
10: qb_estl2:*
11: qb_hlc:*
12: qb_est3:*
13: qb_est4:*
14: qb_est5:*
15: em_estfun:*
16: em_estom:*
17: qb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 184 | 100.0 | 439 | 10 | AM238830 xb34q12.y |
| 2 | 184 | 100.0 | 474 | 12 | BG698550 602658576 |
| 3 | 184 | 100.0 | 491 | 12 | BF724196 bx02a08.y |
| 4 | 184 | 100.0 | 740 | 13 | B1116433 602868855 |
| 5 | 184 | 100.0 | 746 | 13 | B1334702 602998683 |
| 6 | 184 | 100.0 | 768 | 12 | BG687640 602639251 |

| | | | | | |
|----|-------|-------|------|----|----------|
| 7 | 184 | 100.0 | 776 | 13 | B1561125 |
| 8 | 184 | 100.0 | 803 | 13 | B1562083 |
| 9 | 184 | 100.0 | 806 | 13 | B1255914 |
| 10 | 184 | 100.0 | 809 | 12 | BG539839 |
| 11 | 184 | 100.0 | 813 | 12 | BE907823 |
| 12 | 184 | 100.0 | 865 | 12 | BE91785 |
| 13 | 184 | 100.0 | 879 | 14 | BQ216402 |
| 14 | 184 | 100.0 | 882 | 14 | BQ216693 |
| 15 | 184 | 100.0 | 891 | 9 | AL556377 |
| 16 | 184 | 100.0 | 915 | 14 | BQ431952 |
| 17 | 184 | 100.0 | 947 | 9 | AL527764 |
| 18 | 184 | 100.0 | 975 | 14 | BQ055396 |
| 19 | 184 | 100.0 | 1011 | 14 | BQ071045 |
| 20 | 184 | 100.0 | 1026 | 13 | BM557410 |
| 21 | 184 | 100.0 | 1077 | 13 | B1251944 |
| 22 | 184 | 100.0 | 1121 | 14 | BQ067369 |
| 23 | 184 | 100.0 | 1179 | 14 | BQ069475 |
| 24 | 184 | 100.0 | 1310 | 14 | BQ898393 |
| 25 | 184 | 100.0 | 1346 | 14 | BQ071343 |
| 26 | 183 | 99.5 | 366 | 9 | AL561183 |
| 27 | 182.4 | 99.1 | 532 | 10 | AM161050 |
| 28 | 182.4 | 99.1 | 540 | 10 | AM372346 |
| 29 | 182.4 | 99.1 | 715 | 12 | BG721839 |
| 30 | 180.8 | 98.3 | 649 | 14 | BM832648 |
| 31 | 179.2 | 97.4 | 336 | 14 | BM745082 |
| 32 | 177.8 | 96.6 | 630 | 13 | B1826839 |
| 33 | 177.4 | 93.2 | 710 | 13 | B1559395 |
| 34 | 171.2 | 93.0 | 817 | 13 | B1754131 |
| 35 | 171.2 | 93.0 | 973 | 14 | BQ068952 |
| 36 | 171 | 92.9 | 736 | 13 | B1601422 |
| 37 | 168.8 | 91.7 | 625 | 12 | BF529544 |
| 38 | 167.2 | 90.9 | 652 | 13 | B1764233 |
| 39 | 163.2 | 88.7 | 775 | 12 | BG828196 |
| 40 | 160.4 | 87.2 | 869 | 12 | BF571876 |
| 41 | 160 | 87.0 | 592 | 12 | BF978976 |
| 42 | 160 | 87.0 | 687 | 12 | BF791001 |
| 43 | 158.4 | 86.1 | 727 | 12 | BG328041 |
| 44 | 155.8 | 84.7 | 953 | 12 | BF528149 |
| 45 | 154.8 | 84.1 | 905 | 12 | BE792035 |

ALIGNMENTS

RESULT 1
AM238830
LOCUS xb34q12.y1 NCI_CGAP_Lu31 Homo sapiens cDNA clone IMAGE:2578246 5', mRNA sequence.
DEFINITION AM238830 439 bp mRNA linear EST 13-DEC-1999
ACCESSION AM238830
VERSION AM238830.1 GI:6571296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 439)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: www.bio.lnl.gov/dbtp/image/image.html
Seq primer: -40RP from Gibco
High quality sequence stop: 425.
Location/Qualifiers

source

1. .439
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2578246"
/clone_lib="NCI_CGAP_Lu31"
/sex="male"
/dev_stage="Fetal, 14 wk post-conception"
/lab_host="DH10B"
/note="Organ: lung, cell line: Vector: PCMV-SPORT6;
Site_1: EcorV; Site_2: NotI; Cloned unidirectionally, no
5' adaptor. Primer: Oligo dT. Full-length library
constructed by Life Technologies."

BASE COUNT 120 a 85 c 107 g 127 t

Query Match 100.0%; Score 184; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.2e-48;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTCCTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAG 60
|||||
DB 35 ACTTCATTCCTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAG 94
QY 61 CGGTGTTCTCAGAAATTGATAAGACATGCGCAAAACTGTGACATTGGCTTTGAGTA 120
|||||
DB 95 CGGTGTTCTCAGAAATTGATAAGACATGCGCAAAACTGTGACATTGGCTTTGAGTA 154
QY 121 ACCCTGTGCGGTTCTCTATTGACACAAATCAGACCTCAATCCCTAGTAGTAGACA 180
|||||
DB 155 ACCCTGTGCGGTTCTCTATTGACACAAATCAGACCTCAATCCCTAGTAGTAGACA 214
QY 181 TTGA 184
|||||
DB 215 TTGA 218

RESULT 2
BG698550 474 bp mRNA linear EST 07-MAY-2001
LOCUS BG698550
DEFINITION 602658576F2 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4801602 5',
mRNA sequence.
ACCESSION BG698550
VERSION BG698550.1 GI:13965955
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 474)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
Plate: LHAM0694 row: h column: 19
High quality sequence start: 7
High quality sequence stop: 472.
Location/Qualifiers
1. .474
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4801602"
/clone_lib="NCI_CGAP_Skn3"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 123 a 94 c 122 g 135 t

Query Match 100.0%; Score 184; DB 12; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.3e-48;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTCCTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAG 60
|||||
DB 80 ACTTCATTCCTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAG 139
QY 61 CGGTGTTCTCAGAAATTGATAAGACATGCGCAAAACTGTGACATTGGCTTTGAGTA 120
|||||
DB 140 CGGTGTTCTCAGAAATTGATAAGACATGCGCAAAACTGTGACATTGGCTTTGAGTA 199
QY 121 ACCCTGTGCGGTTCTCTATTGACACAAATCAGACCTCAATCCCTAGTAGTAGACA 180
|||||
DB 200 ACCCTGTGCGGTTCTCTATTGACACAAATCAGACCTCAATCCCTAGTAGTAGACA 259
QY 181 TTGA 184
|||||
DB 260 TTGA 263

RESULT 3
BF724196 491 bp mRNA linear EST 05-JAN-2001
LOCUS BF724196
DEFINITION bx02a08.y1 Human Iris cDNA (un-normalized, unamplified): BX Homo
sapiens cDNA clone bx02a08 5', mRNA sequence.
ACCESSION BF724196
VERSION BF724196.1 GI:12040105
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 491)
AUTHORS Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
TITLE NEIBANK: EST analysis and bioinformatics for ocular genomics
JOURNAL Invest. Ophthalmol. Vis. Sci. 41, (2000) In press
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: grame@helix.nih.gov
Plate: 02 row: a column: 08
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1. .491
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="bx02a08"
/clone_lib="Human Iris cDNA (un-normalized, unamplified):
BX"
/tissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMDH10B"
/note="Organ: Eye; Vector: PCMVSPORT6; Post-mortem Iris
tissue was pooled from 10 individuals ranging in age from
4-80 years and RNA was extracted. From this pooled sample
an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A
directionally cloned cDNA library in the PCMVSPORT6 vector
was constructed at Life Technologies, essentially
following the protocols of the Superscript Plasmid System
full details of which are contained in the manufacturer's
instruction manual (<http://www.lifetech.com/>). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-pGACTAGTTCAGATCGGACGGCCGCC(T)15-3'
]. Not I/Dlunt end inserts were cloned into the Not I/Ecor

v sites in the vector. EST analysis was performed on the
unamplified library at the NIH Intramural Sequencing
Center (NISC)."

Query Match 100.0%; Score 184; DB 12; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.3e-48;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BASE COUNT 139 a 95 c 118 g 139 t
ORIGIN
OY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTCTGTTGCTTAACCTTAAGAAG 60
DB 12 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTCTGTTGCTTAACCTTAAGAAG 71
OY 61 CGGTGTTTCTCAGCAATGATAGACCATGCGACAAAACCTGTGACGATTGGCTTGAAGTA 120
DB 72 CGGTGTTTCTCAGCAATGATAGACCATGCGACAAAACCTGTGACGATTGGCTTGAAGTA 131
OY 121 ACCCTGTGCGGTTCCCTATTGTCAGCAAGAAATGACAGCCCTATCCCTTAGTAGGAAGCA 180
DB 132 ACCCTGTGCGGTTCCCTATTGTCAGCAAGAAATGACAGCCCTATCCCTTAGTAGGAAGCA 191

OY 181 TTGA 184
DB 192 TTGA 195

RESULT 4 740 bp mRNA linear EST 26-JUN-2001
LOCUS B116433
DEFINITION 60286855F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5017525 5',
MRNA sequence.

ACCESSION B116433
VERSION B116433.1 GI:14567334
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mhc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: LNCM1829 row: m column: 14
High quality sequence stop: 698.

FEATURES
source
1. 740
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5017525"
/clone_1lb="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2:
ECORI; CDNA made by oligo-dT priming. Directionally
cloned into EORI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 206 a 162 c 199 g 173 t
ORIGIN

Query Match 100.0%; Score 184; DB 13; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.4e-48;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTCTGTTGCTTAACCTTAAGAAG 60
DB 25 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTCTGTTGCTTAACCTTAAGAAG 84
OY 61 CGGTGTTTCTCAGCAATGATAGACCATGCGACAAAACCTGTGACGATTGGCTTGAAGTA 120
DB 85 CGGTGTTTCTCAGCAATGATAGACCATGCGACAAAACCTGTGACGATTGGCTTGAAGTA 144
OY 121 ACCCTGTGCGGTTCCCTATTGTCAGCAAGAAATGACAGCCCTATCCCTTAGTAGGAAGCA 180
DB 145 ACCCTGTGCGGTTCCCTATTGTCAGCAAGAAATGACAGCCCTATCCCTTAGTAGGAAGCA 204
OY 181 TTGA 184
DB 205 TTGA 208

RESULT 5 746 bp mRNA linear EST 30-JUL-2001
LOCUS B1334702
DEFINITION 602998683F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5140628 5',
MRNA sequence.

ACCESSION B1334702
VERSION B1334702.1 GI:15019359
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mhc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11346 row: f column: 21
High quality sequence stop: 744.

FEATURES
source
1. 746
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5140628"
/clone_1lb="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 210 a 158 c 202 g 176 t
ORIGIN
Query Match 100.0%; Score 184; DB 13; Length 746;
Best Local Similarity 100.0%; Pred. No. 1.4e-48;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTCTGTTGCTTAACCTTAAGAAG 60
DB 35 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTCTGTTGCTTAACCTTAAGAAG 94
OY 61 CGGTGTTTCTCAGCAATGATAGACCATGCGACAAAACCTGTGACGATTGGCTTGAAGTA 120
|||||

Db 95 CGGTGTTTCAGAAATGTAAGACCATGACAAAACCTGTGACGATTGCGCTTTGGAGTA 134
 QY 121 ACCCTGTGTCGGGTTCCCTATTGACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAACA 180
 Db 155 ACCCTGTGTCGGGTTCCCTATTGACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAACA 214
 QY 181 TTGA 184
 Db 215 TTGA 218

RESULT 6
 BG687640 768 bp mRNA linear EST 01-MAY-2001
 LOCUS 602639251F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4762356 5',
 DEFINITION mRNA sequence.
 ACCESSION BG687640
 VERSION BG687640.1 GI:13919037
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONETECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LCM1616 row: e column: 13
 High quality sequence stop: 716.
 Location/Qualifiers
 1..768

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4762356"
 /clone_1ib="NIH_MGC_59"
 /tissue_type="mucoepidermoid carcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site_1:
 sfll (ggcgccctggcc); Site_2: sfll (ggccattatggcc);
 Double-stranded cDNA was prepared from cell line RNA. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGCGCCGACATG-dT(30)BN-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
 library."

BASE COUNT 214 a 163 c 212 g 179 t
 ORIGIN

Query Match 100.0%; Score 184; DB 12; Length 768;
 Best Local Similarity 100.0%; Pred. No. 1.5e-48;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTCTTCAGATGTAAGACCATGACAAAACCTGTGACGATTGCGCTTTGGAGTA 60
 Db 58 ACTTCATTCTTCAGATGTAAGACCATGACAAAACCTGTGACGATTGCGCTTTGGAGTA 117
 QY 61 CGGTGTTTCAGAAATGTAAGACCATGACAAAACCTGTGACGATTGCGCTTTGGAGTA 120
 Db 118 CGGTGTTTCAGAAATGTAAGACCATGACAAAACCTGTGACGATTGCGCTTTGGAGTA 177

QY 121 ACCCTGTGTCGGGTTCCCTATTGACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAACA 180
 Db 178 ACCCTGTGTCGGGTTCCCTATTGACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAACA 237
 QY 181 TTGA 184
 Db 238 TTGA 241

RESULT 7
 B1561125 776 bp mRNA linear EST 05-SEP-2001
 LOCUS 603253647F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296306 5',
 DEFINITION mRNA sequence.
 ACCESSION B1561125
 VERSION B1561125.1 GI:15448439
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshitaki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LAM11749 row: m column: 11
 High quality sequence stop: 774.
 Location/Qualifiers
 1..776

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5296306"
 /clone_1ib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (ggcgag
); Oligo-dT primed using primer 5'-TTTCTTTTCTTTTCTT-3',
 size-selected for average insert size 2.2 kb and
 normalized to R01 5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI/NHRI, National
 Institutes of Health). Note: this is a NIH_MGC library."

BASE COUNT 213 a 168 c 216 g 179 t
 ORIGIN

Query Match 100.0%; Score 184; DB 13; Length 776;
 Best Local Similarity 100.0%; Pred. No. 1.5e-48;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTCTTCAGATGTAAGACCATGACAAAACCTGTGACGATTGCGCTTTGGAGTA 60
 Db 65 ACTTCATTCTTCAGATGTAAGACCATGACAAAACCTGTGACGATTGCGCTTTGGAGTA 124
 Db 125 CGGTGTTTCAGAAATGTAAGACCATGACAAAACCTGTGACGATTGCGCTTTGGAGTA 184
 QY 121 ACCCTGTGTCGGGTTCCCTATTGACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAACA 180
 Db 185 ACCCTGTGTCGGGTTCCCTATTGACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAACA 244
 QY 181 TTGA 184

Db 245 TTGA 248
|||||
RESULT 8
LOCUS B1562083
DEFINITION B1562083 803 bp mRNA linear EST 05-SEP-2001
603256666F1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:5298785 5',
mRNA sequence.
ACCESSION B1562083
VERSION B1562083.1 GI:15449409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 803)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11756 row: d column: 18
High quality sequence stop: 766.
Location/Qualifiers
1..803
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5298785"
/clone_1lb="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgaag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT 217 a 181 c 220 g 185 t
ORIGIN
Query Match 100.0%; Score 184; DB 13; Length 803;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTGTCCTGTTGGCTAATTGAAG 60
|||||
Db 83 ACTTCATTTCTTCAGTACAGACAGTGTGTGTCCTGTTGGCTAATTGAAG 142
QY 61 CGGTGTTTTCAGAAATGATTAAGACCATGGCACAAAACGTGTGAGATTGGCTTTGGAGTA 120
|||||
Db 143 CGGTGTTTTCAGAAATGATTAAGACCATGGCACAAAACGTGTGAGATTGGCTTTGGAGTA 202
QY 121 ACCCTGTGTGGCGTTCCTATTGACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGCA 180
|||||
Db 203 ACCCTGTGTGGCGTTCCTATTGACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGCA 262
QY 181 TTGA 184
|||||
Db 263 TTGA 266
RESULT 9

B1255914
LOCUS B1255914 806 bp mRNA linear EST 17-JUL-2001
DEFINITION 602976536F1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:5115705 5',
mRNA sequence.
ACCESSION B1255914
VERSION B1255914.1 GI:14809808
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 806)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11281 row: h column: 10
High quality sequence stop: 779.
Location/Qualifiers
1..806
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5115705"
/clone_1lb="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
BASE COUNT 222 a 178 c 222 g 184 t
ORIGIN
Query Match 100.0%; Score 184; DB 13; Length 806;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTGTCCTGTTGGCTAATTGAAG 60
|||||
Db 60 ACTTCATTTCTTCAGTACAGACAGTGTGTGTCCTGTTGGCTAATTGAAG 119
QY 61 CGGTGTTTTCAGAAATGATTAAGACCATGGCACAAAACGTGTGAGATTGGCTTTGGAGTA 120
|||||
Db 120 CGGTGTTTTCAGAAATGATTAAGACCATGGCACAAAACGTGTGAGATTGGCTTTGGAGTA 179
QY 121 ACCCTGTGTGGCGTTCCTATTGACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGCA 180
|||||
Db 180 ACCCTGTGTGGCGTTCCTATTGACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGCA 239
QY 181 TTGA 184
|||||
Db 240 TTGA 243
RESULT 10
B6339839
LOCUS B6339839 809 bp mRNA linear EST 03-APR-2001
DEFINITION 602563723F1 NIH_MGC_77 Homo sapiens CDNA clone IMAGE:4688452 5',
mRNA sequence.
ACCESSION B6339839
VERSION B6339839.1 GI:13532072
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Query Match 100.0%; Score 184; DB 14; Length 882;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTTCATCTTCAGGTACAGACAGTGTGTGTCTGTGTGGCTAATCTTAAGAAG 60
|||||
DB 60 ACTTCATCTTCAGGTACAGACAGTGTGTGTCTGTGTGGCTAATCTTAAGAAG 119
|||||
OY 61 CGGTGTTCTCAGAAATGTAAGACATGGCACAACAGTGCATTCGCTTGGAGTA 120
|||||
DB 120 CGGTGTTCTCAGAAATGTAAGACATGGCACAACAGTGCATTCGCTTGGAGTA 179
|||||
OY 121 ACCCTGTGTGCGGTTCTTATTCACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGCA 180
|||||
DB 180 ACCCTGTGTGCGGTTCTTATTCACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGCA 239
|||||
OY 181 TTGA 184
|||||
DB 240 TTGA 243

RESULT 15
AL556377 891 bp mRNA linear EST 16-FEB-2001
LOCUS AL556377 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DK004YL02 5
DEFINITION Prime, mRNA sequence.
ACCESSION AL556377
VERSION AL556377.1 GI:12898997
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 891)
AUTHORS Li W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

FEATURES
source
Location/Qualifiers
1..891

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DK004YL02"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 236 a 200 c 251 g 202 t 2 others
ORIGIN

Query Match 100.0%; Score 184; DB 9; Length 891;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTTCATCTTCAGGTACAGACAGTGTGTGTCTGTGTGGCTAATCTTAAGAAG 60
|||||
DB 51 ACTTCATCTTCAGGTACAGACAGTGTGTGTCTGTGTGGCTAATCTTAAGAAG 110
|||||
OY 61 CGGTGTTCTCAGAAATGTAAGACATGGCACAACAGTGCATTCGCTTGGAGTA 120
|||||
DB 111 CGGTGTTCTCAGAAATGTAAGACATGGCACAACAGTGCATTCGCTTGGAGTA 170
|||||

OY 121 ACCCTGTGTGCGGTTCTTATTCACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGCA 180
|||||
DB 171 ACCCTGTGTGCGGTTCTTATTCACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGCA 230
|||||
OY 181 TTGA 184
|||||
DB 231 TTGA 234

Search completed: February 16, 2003, 01:20:13
Job time : 2190.78 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: February 16, 2003, 02:51:59 ; Search time 54.2388 Seconds
(without alignments)
1397.991 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 327
Sequence: 1 acctatcttcagctacag.....ccttagtagtgaagcatlga 184

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+np.model -prv-xlp
-O/cgn2_1/USPTO_spool/US0939293/runat_12022003_170353_8526/app-query.fasta.1.590
-Db=SPREMBL_21 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0939293.ecgn.1.1.51@runat_12022003_170353_8526 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|--------|---------------------|
| 1 | 237 | 72.5 | 157 | 08R1D8 | 08r1d8 mus musculus |

| | | | | | | | |
|---|----|------|------|------|----|--------|---------------------|
| C | 2 | 68 | 21.2 | 212 | 16 | 08ZMN8 | 08zmn8 salmonella |
| C | 3 | 67 | 20.5 | 202 | 10 | 04B630 | 04b630 arabidopsis |
| C | 4 | 66.5 | 20.7 | 949 | 5 | P90956 | P90956 caenorhabdit |
| C | 5 | 66 | 20.6 | 153 | 5 | P91741 | P91741 hydra atten |
| C | 6 | 62.5 | 19.1 | 336 | 10 | 08RW56 | 08rw56 eragrostis |
| C | 7 | 62.5 | 19.1 | 1353 | 10 | 09HF4 | 09hf4 arabidopsis |
| C | 8 | 61.5 | 18.8 | 810 | 11 | 08R5C9 | 08r5c9 mus musculus |
| C | 9 | 61.5 | 18.8 | 2348 | 5 | 09V346 | 09v346 drosophila |
| C | 10 | 61.5 | 18.8 | 2456 | 11 | 070151 | 070151 rattus norv |
| C | 11 | 60.5 | 18.5 | 322 | 10 | 09M0U7 | 09m0u7 arabidopsis |
| C | 12 | 60 | 18.3 | 466 | 16 | 083526 | 083526 treponema p |
| C | 13 | 60 | 18.3 | 1065 | 5 | 09U192 | 09u192 leishmania |
| C | 14 | 59.5 | 18.2 | 233 | 10 | 09LU33 | 09lu33 arabidopsis |
| C | 15 | 59.5 | 18.5 | 332 | 10 | 09ZUN7 | 09zun7 arabidopsis |
| C | 16 | 59 | 18.4 | 417 | 5 | 09VBS9 | 09vbs9 drosophila |
| C | 17 | 59 | 18.0 | 530 | 2 | 093G20 | 093g20 streptomyce |
| C | 18 | 59 | 18.4 | 778 | 4 | 096M37 | 096m37 homo sapien |
| C | 19 | 59 | 18.4 | 1883 | 4 | 09H2Y7 | 09h2y7 homo sapien |
| C | 20 | 58.5 | 18.4 | 351 | 12 | 08UYL0 | 08uy10 wheat dwarf |
| C | 21 | 58.5 | 18.2 | 2233 | 5 | 094711 | 094711 paramedum |
| C | 22 | 58 | 18.1 | 172 | 2 | 08VTK1 | 08vtk1 staphylococ |
| C | 23 | 58 | 18.1 | 401 | 10 | 08S8A2 | 08s8a2 arabidopsis |
| C | 24 | 58 | 17.7 | 687 | 10 | 049728 | 049728 arabidopsis |
| C | 25 | 58 | 18.1 | 943 | 16 | 09ADE2 | 09ade2 streptomyce |
| C | 26 | 58 | 18.1 | 1589 | 5 | 061651 | 061651 drosophila |
| C | 27 | 58 | 18.1 | 1589 | 5 | 001712 | 001712 drosophila |
| C | 28 | 58 | 18.1 | 1589 | 5 | 09V9X8 | 09v9x8 drosophila |
| C | 29 | 57.5 | 17.6 | 394 | 17 | 08Z249 | 08z249 pyrobaculum |
| C | 30 | 57.5 | 17.9 | 468 | 12 | 08Q6G1 | 08q6g1 influenza a |
| C | 31 | 57 | 17.4 | 214 | 2 | 093M22 | 093m22 streptomyce |
| C | 32 | 57 | 17.4 | 233 | 10 | 081513 | 081513 arabidopsis |
| C | 33 | 57 | 17.8 | 258 | 16 | 09L1H7 | 09l1h7 streptomyce |
| C | 34 | 57 | 17.8 | 678 | 5 | 09V306 | 09v306 drosophila |
| C | 35 | 57 | 17.8 | 2589 | 12 | 066776 | 066776 equine rhin |
| C | 36 | 56.5 | 17.3 | 290 | 12 | 091K76 | 091k76 spodoptera |
| C | 37 | 56.5 | 17.6 | 359 | 11 | 08VIE3 | 08vie3 rattus norv |
| C | 38 | 56.5 | 17.3 | 390 | 10 | 09E5Y5 | 09e5y5 cynodon dac |
| C | 39 | 56.5 | 17.3 | 444 | 12 | 09YTK1 | 09ytk1 ateline her |
| C | 40 | 56.5 | 17.6 | 446 | 11 | 08VIE4 | 08vie4 rattus norv |
| C | 41 | 56.5 | 17.6 | 459 | 11 | 08VIE2 | 08vie2 rattus norv |
| C | 42 | 56.5 | 17.6 | 816 | 16 | 0912W1 | 0912w1 pseudomonas |
| C | 43 | 56.5 | 17.6 | 1098 | 2 | 09R2W8 | 09r2w8 borrelia bu |
| C | 44 | 56.5 | 17.6 | 1630 | 13 | 090724 | 090724 gallus gall |
| C | 45 | 56 | 17.1 | 161 | 16 | 09ZL27 | 09zl27 helicobacte |

ALIGNMENTS

RESULT 1

ID 08R1D8 PRELIMINARY; PRT; 157 AA.
AC 08R1D8;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Similar to RIKEN CDNA 0610041G12 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024780; AAH24780.1; -
SQ SEQUENCE 157 AA; 17799 MW; 0F67319F05EAC6E7 CRC64;

Alignment Scores:

Pred. No.: 1.8e-24 Length: 157
Score: 237.00 Matches: 47
Percent Similarity: 85.00% Conservative: 4
Best Local Similarity: 78.33% Mismatches: 7

```

Query Match: 72.48% Indels: 2
DB: 11 Gaps: 1
US-09-939-293-1_COPY_56_239 (1-184) x Q8RLD8 (1-157)

QY 4 TCATTCTTACAGTACAGACAGTGTGTGTCCTGTTGGCTTACTTAAAGACGG 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14 SerLeuPheArgTyrArgInArgPhe-----ProValLeuAlaSerLysArg 31
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 64 TGTTCCTGCAATGTAATGACCATGACAAATGACGATGAGCTTGGAGTAAC 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 32 CysPheSerGluLeuIleLysProTrpHisLysThrValLeuThrGlyPheGlyMetThr 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 124 CTTGTGCGGTTTCCTATTGCACAGAAATCAGAGCTCATTTCCCTTAGTAGTAAGCATTTG 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 52 LeuCySAIaValProIleAlaGlnLysSerGluProGlnSerLeuSerAsnGlnAlaLeu 71

RESULT 2
Q8ZMN8 PRELIMINARY; PRT; 212 AA.
AC Q8ZMN8;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative cytoplasmic protein.
GN STM2766.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_Taxid=602;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2."
RL Nature 413:852-856(2001).
DR EMBL; AB008826; AAL21652.1; -.
KW Hypothetical protein; complete proteome.
SQ SEQUENCE 212 AA; 24460 MW; 697E4049BE7E2C95 CRC64;

Alignment Scores:
Pred. No.: 1.15 Length: 212
Score: 68.00 Matches: 19
Percent Similarity: 43.40% Conservative: 4
Best Local Similarity: 35.85% Mismatches: 26
Query Match: 21.18% Indels: 4
DB: 16 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x Q8ZMN8 (1-212)

QY 149 TTCTGTCAATAGGAACCGACACAGGTTACTCCAAAGCAATCGTCACAGTT----- 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 17 PheCySAIaValArgAspIaPheSerAsnLeuGlnLeuValThrIaGluIle 36
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 95 -----TTGTGCCATGCGTCTTATCATCTTGTGAGAAACCGCTTCTTAAAGTTAGCCACA 42
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 37 LeuAsnThrCysHisGlnSerTrpAsnLysGluThrLysAspPheSerLeuIleSer 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 41 ACAGGAACACACAACACGTGTGTATCGTGAAGATGA 3
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 ThrGlyThrHisLysArgIleLeuValLysThrAsp 69

RESULT 3
Q49630 PRELIMINARY; PRT; 202 AA.
AC Q49630;
DT 01-JUN-1998 (TREMBLrel. 06, Created)

```

```

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 22.7 kDa protein.
GN T10114.80 OR A14G22250.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_Taxid=3702;
RN (1)
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Drost L., Hall C., Hudson S., Ridley P.,
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
RA Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AL021712; CAI6775.1; -.
DR EMBL; AL161556; CAB79180.1; -.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
KW Hypothetical protein; zinc-finger.
SQ SEQUENCE 202 AA; 22711 MW; B994E4AFEB82407F CRC64;

Alignment Scores:
Pred. No.: 1.58 Length: 202
Score: 67.00 Matches: 19
Percent Similarity: 44.23% Conservative: 4
Best Local Similarity: 36.54% Mismatches: 24
Query Match: 20.49% Indels: 5
DB: 10 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x Q49630 (1-202)

QY 11 TCAGTACAGACAGTGTGTGTCCTGTTGGCTTAACTTAAAGACGGTCTTCT 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 148 ThrGlyAsnAspSerValCys-CysValCysMetGlyArgLysGlyAlaAlaPhe 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 71 CAGAATTGATTAAGACATGACGACAAACGTGACGATGCTTGGAGTAAC----- 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 167 eProcySGlyHisThrPheCysArgValCysSerArgIleuLeuTrpLeuAsnArgGlySe 187
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 123 ----CCTGTGTCGGGTTCCCTATTGACACAGAAATC 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 187 rCysProLeuCySAAsnArgProIleIleGluIle 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
P90956 PRELIMINARY; PRT; 949 AA.
AC P90956;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE T01D3.3 protein.
GN T01D3.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN (1)
RP SEQUENCE FROM N.A.
RA Steward C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).

```


| | | | |
|--------|-------|----------|----|
| Score: | 60.50 | Matches: | 14 |
|--------|-------|----------|----|

Percent Similarity: 50.00% Conservative: 4
 Best Local Similarity: 38.89% Mismatches: 15
 Query Match: 18.50% Indels: 3
 DB: 10 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x Q9M0U7 (1-322)

QY 19 AGACAGTGTGTTGTCCTGTTGCTTAACCTTAAGACGGTGT-----TTC 69
 |||||:::|||||
 DB 258 AtgLyCysPheAsnIleuValGlyAsnLeuGlyArgGlyMetLysArgIle 277
 |||||:::|||||
 DB 70 TCAGATTATAGACCATGACCAAACTGTGACGATTGCTTGA 117
 |||||:::|||||
 DB 278 LysGluLeuArgArgPheHisAspSerThrAlaAspTyr-ProPheGly 293
 |||||:::|||||

RESULT 12

083526 PRELIMINARY: PRT: 466 AA.

AC 083526: TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE K+ transport protein (TRKA).
 GN TP0513.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_Taxid=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NICHOLS;
 RC MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardman J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uitterlind T.,
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis
 RT Spirochete.";
 RL Science 281:375-388(1998).
 DR EMBL, AE001227; AAC65501.1; -.
 DR TIGR, TP0513; -.
 DR InterPro: IPR000309; TrKA_Kuplake.
 DR InterPro: IPR003148; TrKA_N.
 DR Pfam: PF02080; TrKA-C; 2.
 DR Pfam: PF02254; TrKA-N; 2.
 KW Complete proteome.
 SO SEQUENCE 466 AA; 49641 MW; 70266ED62B7F49D0 CRC64;

Alignment Scores:
 Pred. No.: 16 Length: 466
 Score: 60.00 Matches: 20
 Percent Similarity: 41.38% Conservative: 4
 Best Local Similarity: 34.48% Mismatches: 18
 Query Match: 18.35% Indels: 16
 DB: 16 Gaps: 3

US-09-939-293-1_COPY_56_239 (1-184) x 083526 (1-466)

QY 28 TTGTGTGTCCT-----GTTGGCTTAACCTTAAGACGGTGT 66
 |||||:::|||||
 DB 210 LeuCyAlaProGluHisMetGlyArgPheTyrGluLeuAlaGlyPheLys----- 226
 |||||:::|||||
 QY 67 TTCCTCAGATTATAGACCATGACCAAACTGTGACGATTGCTTGA 120
 |||||:::|||||
 DB 227 -----IleHisProValLysLysIleAlaLeuIleGlyMetSerAlaValGly 242
 |||||:::|||||
 QY 121 ACCCTGTGTCGGTTCCTATGTCACAGAATCAGAGCTCATCCCTTAGT 174
 |||||:::|||||
 DB 243 ThrLeuValAlaGlnAspValAlaGluLysCysLysProHisPhePheSer 260
 |||||:::|||||
 RESULT 13

Q9U192 PRELIMINARY: PRT: 1065 AA.

ID Q9U192: TREMBLrel. 13, Created)
 AC Q9U192: (1-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Probable protein kinase.
 GN L2903.06.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_Taxid=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Hilbert H., Wedler H., Wedler E., Duesterhoeft A., Ivens A.C.,
 RA Lawson D., Quail M., Rajandream M.A., Barrell B.G.;
 RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RX MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome.";
 RL Genome Res. 8:135-145(1998).
 DR EMBL, AL117319; CAB5520.1; -.
 DR InterPro: IPR000719; Euk.pkinase.
 DR Pfam: PF00069; pkinase; 2.
 DR PROSITE: PS00001; Euk.pkinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Kinase; Transferase.
 SO SEQUENCE 1065 AA; 112308 MW; 8693736DA18F6BF4 CRC64;

Alignment Scores:
 Pred. No.: 16.8 Length: 1065
 Score: 60.00 Matches: 21
 Percent Similarity: 42.19% Conservative: 6
 Best Local Similarity: 32.81% Mismatches: 19
 Query Match: 18.35% Indels: 18
 DB: 5 Gaps: 4

US-09-939-293-1_COPY_56_239 (1-184) x Q9U192 (1-1065)

QY 19 AGACAGTGTGTTGTCCTGTTGCTTAACCTTAAGACGGTGTTCGAAATTG 78
 ||| |||||:::|||||
 DB 1006 ArgSerCysLeuAspProAla-----GluArgArgThrValPheGluLeu 1022
 |||||:::|||||
 QY 79 ATAGA---CCATGGCACAAAACGTGACGATTGCTTGGAGTAACCTG----- 126
 |||||:::|||||
 DB 1023 PheArgHisProTyr-----IleArgGlyGlyGluGlyAlaThrLeuValGluAla 1039
 |||||:::|||||
 QY 127 -----TGTGGGTTCTATGTCACAGAAATCAGAGCTCAT 162
 |||||:::|||||
 DB 1040 AspThrSerAlaGlySerGlnArgCysSerProProAlaAlaGluLysGlyAlaAspVal 1059
 |||||:::|||||
 QY 163 TCCCTTAGTAGT 174
 |||||:::|||||
 DB 1060 GluLeuSerSer 1063
 |||||:::|||||

RESULT 14

Q9U192 PRELIMINARY: PRT: 233 AA.

ID Q9U192: TREMBLrel. 15, Created)
 AC Q9U192: (1-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE Genomic DNA, chromosome 3, pl clone: MX18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_Taxid=3702;

```

RN [1] SEQUENCE FROM N.A.
RP STRAIN-COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
DR EMBL: AB023045; BAB01718.1;
SQ SEQUENCE 233 AA; 26517 MW; 9A795F53A6C96F51 CRC64;

Alignment Scores:
Pred. No.: 18.1 Length: 233
Score: 59.50 Matches: 14
Percent Similarity: 47.83% Conservative: 8
Best Local Similarity: 30.43% Mismatches: 15
Query Match: 18.208 Indels: 9
DB: 10 Gaps: 2

US-09-939-293-1_COPY_56_239 (1-184) x 09LU33 (1-233)
QY 58 AAGCGGCTTTCAGAAATGATTAAGACCATGGCACAACAACTGTC---ACGATTTGGCTTT 114
Db 154 Gtldtucystyrtsmsneuthrlltelysprottrphtslgltprlleserterlaalaale 173
QY 115 GGATGACCCCTGTGTGGCGTTCCT-----ATTGCACAGAAA 150
Db 174 LysValAlaLeuLysLeuValPProAsnAsnAsnThrPheIleAsnValLeuAlaLalys 193
QY 151 TCAGAGCCCTCATTCCTT 168
Db 194 AspGluThrHisGlnMet 199

RESULT 15
09ZUN7 PRELIMINARY: PRT: 332 AA.
AC 09ZUN7;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative esterase.
GN AT2G19550.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eumetazoa; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Breill C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanhaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen S.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RT Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN EMBL: AC005917; AAD10154.1;

```

```

DR InterPro: IPR000379; Set_estri_sle-
SQ SEQUENCE 332 AA; 37964 MW; AE16C973A56D441 CRC64;

Alignment Scores:
Pred. No.:      18.5          Length:      332
Score:          59.50         Matches:     13
Percent Similarity: 50.00%    Conservative: 6
Best Local Similarity: 34.21%  Mismatches: 18
Query Match:    18.34%       Indels:      1
DB:             10           Gaps:        1

US-09-939-293_1_COPY_56_239 (1-184) x Q9ZUN7 (1-332)

QY 140 ATAGGAACCCACACAGGGTTTACTCAAAAGCCAATCGTCACAGTTTGTCACATGGCTT 81
   ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 17 ValGIgLeuLeuMIsGluThrGLySerIysGLuValVal--ValLeuCYSHISGLYPHe 35
   : 80 ATCAATTCTGTGAGAAMACCAGCGCTTTCTTAAGTTAGGCCACACAGCACACACA 27
   :: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 36 ArgserAspLysThrasmnYSLleuLYsnValatIrntAlaleuGLuLyS 53

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 16, 2003, 01:22:19 : Search time 17.5075 Seconds
(without alignments)
871.816 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 327
Sequence: 1 actcatctcttcagtgacag.....ccttagtagtaagaattga 184

Scoring table:

| | |
|---------------------------|--|
| BLOSUM62 | |
| Xgapop 10.0 , Xgapext 0.5 | |
| Ygapop 10.0 , Ygapext 0.5 | |
| Fgapop 6.0 , Fgapext 7.0 | |
| Delop 6.0 , Delext 7.0 | |

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlp
-Q/cqn2_1/USPTO.spool/US09939293/runat_12022003.170353.8513/app_query.fasta_1.590
-DB=SwissProt_40 -OPMT=fastan -SUFFIX=isp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCAIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09939293.6CCN_1_1_10.8runat_12022003.170353.8513 -NCP=6 -ICPU=3
-NO_XLPEXT -NO_MMMP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120
-MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 327 | 100.0 | 239 | 1 | SMAC_HUMAN |
| 2 | 245 | 74.9 | 237 | 1 | SMAC_MOUSE |
| 3 | 65.5 | 20.0 | 2346 | 1 | COAL_RAT |
| 4 | 65.5 | 20.0 | 2346 | 1 | COAL_HUMAN |
| 5 | 64.5 | 19.7 | 2346 | 1 | COAL_BOVIN |
| 6 | 64.5 | 19.7 | 2346 | 1 | COAL_SHEEP |
| 7 | 61.5 | 18.8 | 2483 | 1 | COAL_HUMAN |
| 8 | 60.5 | 18.5 | 2324 | 1 | COAC_CHICK |
| 9 | 58.5 | 17.9 | 902 | 1 | SYGL_YEAST |
| 10 | 57 | 17.8 | 297 | 1 | FTIR_METH |
| 11 | 57 | 17.4 | 555 | 1 | MASV_PICAN |
| 12 | 56.5 | 17.6 | 446 | 1 | CNTB_MOUSE |
| 13 | 56 | 17.4 | 1301 | 1 | SAC3_YEAST |
| 14 | 55.5 | 17.0 | 310 | 1 | VU04_HSV7J |
| 15 | 55.5 | 17.0 | 449 | 1 | TL04_SPIOL |
| 16 | 55.5 | 17.0 | 830 | 1 | LEM3_HUMAN |
| 17 | 55.5 | 17.0 | 3133 | 1 | HMC3_BOMO |
| 18 | 55 | 16.8 | 64 | 1 | SCX8_MESMA |

| | | | | | | |
|----|------|------|------|---|-------------|--------------------|
| 19 | 55 | 16.8 | 1798 | 1 | LMB2_HUMAN | P55268 homo sapien |
| 20 | 54.5 | 16.7 | 746 | 1 | EXT1_CRIGR | 091882 cricetus |
| 21 | 54.5 | 16.7 | 746 | 1 | EXT1_HUMAN | 016394 homo sapien |
| 22 | 54.5 | 16.7 | 746 | 1 | EXT1_MOUSE | P97464 mus musculu |
| 23 | 54.5 | 17.0 | 941 | 1 | VDP_MOUSE | 092160 mus musculu |
| 24 | 54.5 | 17.0 | 1786 | 1 | YCF1_ARATH | P56785 arabidopsis |
| 25 | 54.5 | 17.0 | 2594 | 1 | 7LES_DROVI | P20806 drosophila |
| 26 | 54 | 16.8 | 354 | 1 | VPRP_HSV7J | P52440 human herpe |
| 27 | 54 | 16.8 | 716 | 1 | RRP2_IAMWL | P15659 influenza a |
| 28 | 54 | 16.5 | 2067 | 1 | BIMB_EMENT | P33144 emeritella |
| 29 | 53.5 | 16.4 | 110 | 1 | Y100_YEAST | P40584 saccharomyc |
| 30 | 53.5 | 16.4 | 317 | 1 | ISPE_ANASP | 089561 anabaena sp |
| 31 | 53.5 | 16.7 | 450 | 1 | CNTB_HUMAN | 09056 homo sapien |
| 32 | 53.5 | 16.4 | 509 | 1 | C982_SOYBN | 048922 glycine max |
| 33 | 53.5 | 16.4 | 615 | 1 | UGST_WHEAT | P27736 triticum ae |
| 34 | 53.5 | 16.4 | 616 | 1 | NGF1_COTJA | P5392 stylosanthe |
| 35 | 53.5 | 16.7 | 662 | 1 | SURT2_STYHA | P53392 stylosanthe |
| 36 | 53.5 | 16.7 | 667 | 1 | SURT1_STYHA | P53391 stylosanthe |
| 37 | 53.5 | 16.7 | 1901 | 1 | YCF1_TOBAC | P12222 nicotiana t |
| 38 | 53 | 16.5 | 276 | 1 | MUR1_LACBR | P48797 lactobacill |
| 39 | 53 | 16.2 | 405 | 1 | CPXM_BACSU | P27632 bacillus su |
| 40 | 53 | 16.5 | 716 | 1 | RRP2_IAMLO | P13168 influenza a |
| 41 | 53 | 16.5 | 716 | 1 | RRP2_IAMTE | P13169 influenza a |
| 42 | 53 | 16.2 | 1184 | 1 | FBL2_HUMAN | P98095 homo sapien |
| 43 | 53 | 16.2 | 1221 | 1 | FBL2_MOUSE | P37889 mus musculu |
| 44 | 53 | 16.2 | 1408 | 1 | SERR_DROME | P18168 drosophila |
| 45 | 53 | 16.2 | 2907 | 1 | FBN2_MOUSE | G61555 mus musculu |

ALIGNMENTS

| | | | | |
|----------|--|---------------------------------|------|---------|
| RESULT 1 | SMAC_HUMAN | STANDARD: | PRT: | 239 AA. |
| ID | SMAC_HUMAN | Q9NR28; Q9NR11; Q9NAV6; Q96LV0; | | |
| AC | Q9NR28; Q9NR11; Q9NAV6; Q96LV0; | | | |
| DT | 16-OCT-2001 (Rel. 40, Created) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update) | | | |
| DE | 15-JUN-2002 (Rel. 41, Last annotation update) | | | |
| DE | Smac protein, mitochondrial precursor (Second mitochondria-derived | | | |
| DE | activator of caspase) (Direct IAP binding protein with low pI). | | | |
| GN | SMAC OR DIABLO. | | | |
| OS | Homo sapiens (human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1), PARTIAL SEQUENCE, FUNCTION, AND TISSUE | | | |
| RP | SPECIFICITY. | | | |
| RP | MEDLINE=20383536; PubMed=10929711; | | | |
| RX | MEDLINE=20383536; PubMed=10929711; | | | |
| RA | Du C., Fang M., Li Y., Li L., Wang X.; | | | |
| RT | "Smac, a mitochondrial protein that promotes cytochrome c-dependent | | | |
| RT | caspase activation by eliminating IAP inhibition." | | | |
| RL | Cell 102:33-42(2000). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1). | | | |
| RP | Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T., | | | |
| RA | Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T., | | | |
| RA | Nakanura Y., Isogai T., Sugano S.; | | | |
| RT | "NEO human cDNA sequencing project." | | | |
| RL | Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases. | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION. | | | |
| RP | PubMed=10950947; | | | |
| RA | Srinivasula S.M., Datta P., Fan X.J., Fernandes-Alnemri T., Huang Z., | | | |
| RA | Alnemri E.S.; | | | |
| RT | "Molecular determinants of the caspase-promoting activity of | | | |
| RT | Smac/DIABLO and its role in the death receptor pathway." | | | |
| RL | J. Biol. Chem. 275:36152-36157(2000). | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1). | | | |
| RP | TISSUE=Cerebellum; | | | |
| RC | Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y., | | | |
| RA | Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M., | | | |

RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
 RA Kawakami B., Nagai K., Isogai T., Sugano S.,
 RT "NEBO human cDNA sequencing project."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOPFORM 1).
 RC TISSUE=Muscle, and uterus;
 RL Strausberg R.,
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 56-239.
 RN Chai J., Du C., Wu J.W., Kyin S., Wang X., Shi Y.,
 RA "Structural and biochemical basis of apoptotic activation by
 RT Smac/DIABLO."
 RL Nature 406:855-862(2000).
 RN [7]
 RP STRUCTURE BY NMR OF 56-64 IN COMPLEX WITH BIRC4.
 RX MEDLINE-21020961; PubMed-11140637;
 RA Liu Z., Sun C., Olejniczak E.T., Meadows R.P., Betz S.F., Oost T.,
 RA Hermann J., Wu J.C., Pesik S.W.,
 RT "Structural basis for binding of Smac/DIABLO to the XIAP BIR3
 RT domain."
 RL Nature 408:1004-1008(2000).
 CC -1- FUNCTION: PROMOTES APOPTOSIS BY ACTIVATING CASPASES IN THE
 CC CYTOCHROME C/PAF-1/CASPASE-9 PATHWAY. ACTS BY OPPOSING THE
 CC INHIBITORY ACTIVITY OF INHIBITOR OF APOPTOSIS PROTEINS (IAP).
 CC -1- SUBUNIT: Homodimer. Interacts with BIRC2, BIRC3, BIRC4/XIAP and
 CC BIRC7.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL BUT RELEASED INTO THE CYTOSOL
 CC WHEN CELLS UNDERGO APOPTOSIS.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2/DIABLO-S;
 CC are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: UBICUITOUSLY EXPRESSED WITH HIGHEST EXPRESSION
 CC IN TESTIS. EXPRESSION IS ALSO HIGH IN HEART, LIVER, KIDNEY,
 CC SPLEEN, PROSTATE AND OVARY. LOW IN BRAIN, LUNG, THYMUS AND
 CC PERIPHERAL BLOOD LEUKOCYTES.
 CC -1- DOMAIN: The mature N-terminus mediates interaction with
 CC BIRC4/XIAP.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF262240; AAF87716.1; -
 DR EMBL: AK024768; BAB14994.1; -
 DR EMBL: AF298770; AAG22077.1; -
 DR EMBL: AK057778; BAB71568.1; -
 DR EMBL: BC004417; AAH04417.1; -
 DR PDB: 1FEW; 13-SEP-00.
 DR PDB: 1G3F; 10-JAN-01.
 DR MIM: 605219; -
 KW Transist peptide; Mitochondrion; Apoptosis; Alternative splicing;
 KW 3D-structure.
 FT TRANSIT 1 55 MITOCHONDRION.
 FT CHAIN 56 239 SMAC PROTEIN.
 FT SITE 56 60 IAP-BINDING MOTIF (BY SIMILARITY).
 FT VARSPLIC 1 60 MAALSKMSLSRYSYTFRRROCCVAVANFKRCSLIRP
 FT WHKVTIGEGVTLCAVPIA -> MKSDYF (IN
 FT ISOFORM 2).
 FT CONFLICT 32 32 K -> E (IN REF. 4).
 FT CONFLICT 44 44 K -> R (IN REF. 2).
 FT CONFLICT 62 105 MISSING (IN REF. 4).
 FT CONFLICT 165 165 E -> K (IN REF. 4).
 SO SEQUENCE 239 AA; 27131 MM; 70C2AE0DC654D031 CPG64;

Alignment Scores: 5.38e-36 Length: 239
 Pred. No.: 327.00 Matches: 61
 Score:

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0
 US-09-939-293-1_COPY_56_239 (1-184) x SMAC_HUMAN (1-239)
 QY 1 ACTTCATTCTTCAGGTACAGACAGTGTGTGTCTCTGTGTGCGTAACCTTAAGAAG 60
 DB 13 ThrSerPhePheArgTyrArgGlnCysLeuValProValValAlaAsnPhelysLys 32
 QY 61 CGGTGTTCTCAGATTGTGAAGACCATGGCCAAACATGTGACATTGGCTTGGAGTA 120
 DB 33 ArgCysPheSerGlnLeuIleArgProPheLysTrpValThrIleGlyPheGlyVal 52
 QY 121 ACCCTGTGTGGCGTCTCTATTGCACAGAAATCAGACGCTCATTCCTTAGTGAAGCA 180
 DB 53 ThrLeuCysAlaValProIleAlaGlnLysSerGlnProHisSerLeuSerSerGluAla 72
 QY 181 TTG 183
 DB 73 Leu 73
 RESULT 2
 SMAC_MOUSE STANDARD; PRT; 237 AA.
 ID Q9JIO3; Q9CZD1; Q9DCD3;
 AC Q9JIO3; Q9CZD1; Q9DCD3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Smac protein, mitochondrial precursor (second mitochondria-derived
 DE activator of caspase) (Direct IAP binding protein with low pI).
 GN SMAC OR DIABLO.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
 RP SPECIFICITY.
 RC STRAIN=BA1B/G; TISSUE=Kidney;
 RX MEDLINE-20383537; PubMed-10929712;
 RA Verhagen A.M., Ekert P.G., Pakusch M., Silve J., Connolly L.M.,
 RA Reid G.E., Moritz R.L., Simpson R.J., Vaux D.L.,
 RT "Identification of DIABLO, a mammalian protein that promotes apoptosis
 RT by binding to and antagonizing IAP proteins."
 RL Cell 102:43-53(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE-21085660; PubMed-11217851;
 RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann S., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baird G.,
 RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Bustinleim M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C.,
 RA Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H.,
 RA Kohlsaki S.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: PROMOTES APOPTOSIS BY ACTIVATING CASPASES IN THE
 CC CYTOCHROME C/PAF-1/CASPASE-9 PATHWAY. ACTS BY OPPOSING THE


```
DR Pfam: PF00289; CPSase-L-chain; 1.
DR Pfam: PF00364; biotin_lipoyl; 1.
DR Pfam: PF01039; Carboxyl_trans; 1.
DR Pfam: PF02785; biotin_carb_C; 1.
DR Pfam: PF02786; CPSase-L.D2; 1.
DR PROSITE: PS00188; BIOTIN; 1.
DR PROSITE: PS00866; CPSASE_1; 1.
DR PROSITE: PS00867; CPSASE_2; 1.
DR Fatty acid biosynthesis; Biotin; Ligase; Multifunctional enzyme;
KW ATP-binding; Phosphorylation.
FT NP_BIND 314 319 ATP (POTENTIAL).
FT ACT_SITE 440 440 BY SIMILARITY.
FT BINDING 785 785 BIOTIN.
FT MOD_RES 77 77 BIOTIN.
FT MOD_RES 79 79 PHOSPHORYLATION.
FT MOD_RES 1200 1200 PHOSPHORYLATION.
FT MOD_RES 1958 1987 COENZYME A-BINDING (BY SIMILARITY).
SQ SEQUENCE 2345 AA; 265191 MW; 78E9CF9ADE1E8771 CRC64;

Alignment Scores:
Pred. No.: 1.17 Length: 2345
Score: 65.50 Matches: 15
Percent Similarity: 51.02% Conservative: 10
Best Local Similarity: 30.61% Mismatches: 13
Query Match: 20.03% Indels: 11
DB: Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x COAL_RAT (1-2345)

QY 67 TTCTCAGAAATTGATAGACCATGCGACAAACTGTGACGATTGGC----- 111
DB 1959 PhesergluilewetiInProtrpalaGlnThValaValaGlyArglaArgleugly 1978
QY 112 -----TTTGAGTAACCCGTGTGCGGTCCTTCATGACAGGAATCA 153
DB 1979 GlylleProvalaGlyValaValaValaGluThrArgThrValaGluLeuSerValProala 1998
QY 154 GAGCCTCATTCCTTAGTAGTGAACCA 180
DB 1999 AspproAlaAsnLeuAspserGluAla 2007

RESULT 4
COAL_HUMAN
ID COAL_HUMAN STANDARD; PRT; 2346 AA.
AC Q13085;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acetyl-CoA carboxylase 1 (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin
  carboxylase (EC 6.3.4.14)].
GN ACACA OR ACAC OR ACC1 OR ACCA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=95249602; PubMed=7732023;
RX Abu-Elheiga L., Jayakumar A., Baldini A., Chitrula S.S., Wakil S.J.,
  "Human acetyl-CoA carboxylase: characterization, molecular cloning,
  RT and evidence for two isoforms."
  Proc. Natl. Acad. Sci. U.S.A. 92:4011-4015(1995).
CC -!- FUNCTION: CATALYZES THE RATE-LIMITING REACTION IN THE BIOGENESIS
  OF LONG-CHAIN FATTY ACIDS. THIS PROTEIN CARRIES THREE FUNCTIONS:
  CC BIOTIN CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND
  CC CARBOXYLTRANSFERASE.
CC -!- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-) = ADP + phosphate
  CC + malonyl-CoA.
CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
  CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -!- COFACTOR: BIOTIN.
CC -!- ENZYME REGULATION: BY PHOSPHORYLATION (BY SIMILARITY).
```

```

DE carboxylase (EC 6.3.4.14)].
OS ACACA OR ACAC OR ACCA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OX Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RC MEDLINE=21378179; PubMed=11485560;
RA Mao J., Marcos S., Davis S.K., Buzalaf J., Seyfert H.M.;
RT "Genomic distribution of three promoters of the bovine gene encoding
RT acetyl-CoA carboxylase alpha and evidence that the nutritionally
RT regulated promoter 1 contains a repressive element different from
RT that in rat.";
RL Biochem. J. 358:127-135(2001).
CC -1- FUNCTION: CATALYZES THE RATE-LIMITING REACTION IN THE BIOGENESIS
CC OF LONG-CHAIN FATTY ACIDS. THIS PROTEIN CARRIES THREE FUNCTIONS:
CC BIOTIN CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND
CC CARBOXYLTRANSFERASE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-) -> ADP + phosphate
CC + malonyl-CoA.
CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -1- COFACTOR: BIOTIN.
CC -1- ENZYME REGULATION: BY PHOSPHORYLATION (BY SIMILARITY).
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first (rate limiting)
CC step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AJ132890; CAB56826.1; -.
CC DR HSSP; P24182; IDVL.
CC DR InterPro: IPR001882; Biotin_attach.
CC DR InterPro: IPR000089; Biotin_lipoyl.
CC DR InterPro: IPR000901; CPSase.
CC DR InterPro: IPR000022; Carboxyl_trans.
CC DR Pfam; PF00289; CPSase_1; 1.
CC DR Pfam; PF00364; biotin_lipoyl; 1.
CC DR Pfam; PF01039; Carboxyl_trans; 1.
CC DR Pfam; PF02785; Biotin_carb_C; 1.
CC DR Pfam; PF02786; CPSase_L_D2; 1.
CC DR PROSITE; PS00188; BIOTIN; 1.
CC DR PROSITE; PS00866; CPSASE_1; 1.
CC DR PROSITE; PS00867; CPSASE_2; 1.
CC KM Fatty acid biosynthesis; Biotin; lysase; Multifunctional enzyme;
CC ATP-binding; Phosphorylation.
FT NP_BIND 315 320 ATP (POTENTIAL).
FT ACT_SITE 441 441 BY SIMILARITY.
FT BINDING 786 786 BIOTIN (BY SIMILARITY).
FT DOMAIN 1959 1988 COENZYME A-BINDING (BY SIMILARITY).
FT MOD_RES 78 78 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 80 80 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 1201 1201 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 2346 AA; 265301 MW; 32886C5D03EEAE0E CRC64;

```

Alignment Scores:

```

Pred. No.: 1 6
Score: 64.50
Percent Similarity: 51.028
Best Local Similarity: 30.618
Query Match: 19.728
DB: 1

```

```

Length: 2346
Matches: 15
Conservative: 10
Mismatch: 13
Indels: 11
Gaps: 1

```

```

US-09-939-293-1_COPY_56_239 (1-184) x COAL_BOVIN (1-2346)
QY 67 TTCTCAGATGTGATAGACCATGCGACAAAACCTGTGACATTGCG----- 111
DB 1960 PheSerGluIleMetGlnProTPalaGlnThValValGlyIyrGalaArgLeuGly 1979
QY 112 -----TTGGATGTAACCCCTGTGCGGTTCTCTATTGACAGAAATCA 153
DB 1980 GlyIleProValGlyValAlaValAlaGluThrArgThValGluIleuSerIleProAla 1999
QY 154 GAGCCTCATYCCCTAGTAGTAAGCA 180
DB 2000 AspProAlaAsnLeuSpsSerGuaAla 2008
RESULT 6
COAL_SHEEP
ID COAL_SHEEP STANDARD; PRT; 2346 AA.
AC Q28559;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acetyl-CoA carboxylase 1 (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin
DE carboxylase (EC 6.3.4.14)].
GN ACACA OR ACAC.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Finn-Dorset; TISSUE=adipose tissue;
RX MEDLINE=95197015; PubMed=7890176;
RA Barber M.C., Travers M.T.;
RT "Cloning and characterisation of multiple acetyl-CoA carboxylase
RT transcripts in ovine adipose tissue.";
RL Gene 154:271-275(1995).
CC -1- FUNCTION: CATALYZES THE RATE-LIMITING REACTION IN THE BIOGENESIS
CC OF LONG-CHAIN FATTY ACIDS. THIS PROTEIN CARRIES THREE FUNCTIONS:
CC BIOTIN CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND
CC CARBOXYLTRANSFERASE.
CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-) -> ADP + phosphate
CC + malonyl-CoA.
CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -1- COFACTOR: BIOTIN.
CC -1- ENZYME REGULATION: BY PHOSPHORYLATION (BY SIMILARITY).
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first (rate limiting)
CC step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X80045; CA56352.1; -.
CC DR HSSP; P24182; IDVL.
CC DR InterPro: IPR001882; Biotin_attach.
CC DR InterPro: IPR000089; Biotin_lipoyl.
CC DR InterPro: IPR000901; CPSase.
CC DR InterPro: IPR000022; Carboxyl_trans.
CC DR Pfam; PF00289; CPSase_L_chain; 1.
CC DR Pfam; PF00364; biotin_lipoyl; 1.
CC DR Pfam; PF01039; Carboxyl_trans; 1.
CC DR Pfam; PF02785; Biotin_carb_C; 1.
CC DR Pfam; PF02786; CPSase_L_D2; 1.
CC DR PROSITE; PS00188; BIOTIN; 1.
CC DR PROSITE; PS00866; CPSASE_1; 1.

```



```

Db      2111 ArgAlaArgLeuGlyIleProValGlyValIleAlaValGluThrArgThrValGlu 2130
OY      139  ATTCACAGAAATCAGAGCCTCATTCCTTACTAGTAGGAACA 180
Db      2131 ValAlaValProAlaAspProAlaAsnLeuAspSerGluAla 2144

RESULT 8
COAC.CHICK STANDARD; PRT; 2324 AA.
AC      P11029;
DT      01-JUL-1989 (Rel. 11, Created)
DT      01-JUL-1989 (Rel. 11, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Acetyl-CoA carboxylase (EC 6.4.1.2) (ACC) [includes: Biotin
GN      carboxylase (EC 6.3.4.14)].
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC      TISSUE=Liver;
RX      MEDLINE=88139305; PubMed=2893793;
RA      Takai T., Yokoyama C., Wada K., Tanabe T.;
RT      "Primary structure of chicken liver acetyl-CoA carboxylase deduced
RT      from cDNA sequence."
RL      J. Biol. Chem. 263:2651-2657(1988).
RN      [2]
RP      SEQUENCE OF 493-820 FROM N.A.
RC      TISSUE=Liver;
RX      MEDLINE=87106011; PubMed=2879745;
RA      Takai T., Wada K., Tanabe T.;
RT      "Primary structure of the biotin-binding site of chicken liver
RT      acetyl-CoA carboxylase."
RL      FEBS Lett. 212:98-103(1987).
CC      -1- FUNCTION: CATALYZES THE RATE-LIMITING REACTION IN THE BIOGENESIS
CC      OF LONG-CHAIN FATTY ACIDS. THIS PROTEIN CARRIES THREE FUNCTIONS:
CC      BIOTIN CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND
CC      CARBOXYLTRANSFERASE.
CC      -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-) = ADP + phosphate
CC      + malonyl-CoA.
CC      -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC      = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC      -1- COFACITOR: BIOTIN.
CC      -1- ENZYME REGULATION: BY PHOSPHORYLATION.
CC      -1- PATHWAY: Long-chain fatty acid biosynthesis; first (rate limiting)
CC      step.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; J03541; AAA48701.1; -
DR      EMBL; X05019; CAA28675.1; -
DR      PIR; A27903; A27903.
DR      PIR; A29337; A29337.
DR      PIR; A29924; A29924.
DR      HSSP; p24182; 1DVI.
DR      InterPro; IPR001882; Biotin_attach.
DR      InterPro; IPR000089; Biotin_lipoyl.
DR      InterPro; IPR000901; CPSase.
DR      InterPro; IPR000022; Carboxyl_trans.
DR      Pfam; PF00289; CPSase_L_chain; 1.
DR      Pfam; PF00364; biotin_lipoyl; 1.

```

```

DR      Pfam; PF01039; Carboxyl_trans; 1.
DR      Pfam; PF02785; Biotin_carb_C; 1.
DR      Pfam; PF02786; CPSase_L_D2; 1.
DR      PROSITE; PS00188; BIOTIN; 1.
DR      PROSITE; PS00866; CPSASE_1; 1.
DR      PROSITE; PS00867; CPSASE_2; 1.
KW      Fatty acid biosynthesis; Biotin; Ligase; Multifunctional enzyme;
KW      ATP-binding; Phosphorylation.
FT      NP_BIND 315 320
FT      ACT_SITE 441 441
FT      BINDING 786 786
FT      MOD_RES 78 78
FT      MOD_RES 80 80
FT      MOD_RES 1193 1193
FT      DOMAIN 1936 1965
SQ      SEQUENCE 2324 AA; 262717 MW; 3F1C541F01BBE6 CRC64;

Alignment Scores:
Pred. No.: 5.54 Length: 2324
Score: 60.50 Matches: 15
Percent Similarity: 46.30% Conservative: 10
Best Local Similarity: 27.78% Mismatches: 18
Query Match: 18.50% Indels: 11
DB: Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x COAC.CHICK (1-2324)
OY      52  TTTAAGAGCGGTGTTTCAGATTGATTAAGACCATGGCAAAACGTGACGATTGCC 111
Db      1932  PheaSpasnclySerPheLeuGluIleMetGlnProTyrPheGlnThrValValGly 1951
OY      112  -----TTTGAGTAAACCTGTCGTGGCTTCCT 138
Db      1952  ArgAlaArgLeuGlyIleProValGlyValIleAlaValGluThrArgThrValGlu 1971
OY      139  ATTCACAGAAATCAGAGCCTCATTCCTTACTAGTAGGAACA 180
Db      1972  LeuSerIleProAlaAspProAlaAsnLeuAspSerGluAla 1985

RESULT 9
ID      SYG1_YEAST STANDARD; PRT; 902 AA.
AC      P40528; P40964;
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      01-NOV-1995 (Rel. 32, Last annotation update)
DE      Syg1 protein.
GN      SYG1 OR YII047C.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX      NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SP1;
RA      Spain B.H., Koo D., Ramakrishnan M., Dzugzor B., Colicelli J.;
RL      Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=S288c / AB972;
RA      Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA      Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA      Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA      Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA      Rajandream M.A., Riles L., Riles T., Rowley N., Skelton J., Smith V.,
RA      Walsh S.V., Whitehead S.;
RL      Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; Z46861; CAAB6904.1; -
DR EMBL; U14726; AAA91621.1; -
DR PIR; S48245; S48245.
DR SGD; S0001309; SYG1.
DR InterPro; IPR004342; EXS_Cterm.
DR InterPro; IPR004331; SPX.
DR Pfam; PF03105; SPX; 1.
DR Pfam; PF03124; EXS; 1.
KW Transmembrane.
FT TRANSMEM 405 425 POTENTIAL.
FT TRANSMEM 436 456 POTENTIAL.
FT TRANSMEM 498 518 POTENTIAL.
FT TRANSMEM 524 544 POTENTIAL.
FT TRANSMEM 555 575 POTENTIAL.
FT TRANSMEM 576 596 POTENTIAL.
FT TRANSMEM 674 694 POTENTIAL.
FT TRANSMEM 733 753 POTENTIAL.
FT TRANSMEM 762 782 POTENTIAL.
FT CONFLICT 177 177 S->T (IN REF. 1).
SQ SEQUENCE 902 AA; 104217 MW; F8D87D1DB3AED64 CRC64;

Alignment Scores:
Pred. No.: 9.66 Length: 902
Score: 58.50 Matches: 16
Percent Similarity: 45.61% Conservative: 10
Best Local Similarity: 28.07% Mismatches: 16
Query Match: 17.89% Indels: 15
Gaps: 2

US-09-939-293-1_COPY_56_239 (1-184) x SYG1_YEAST (1-902)

OY 4 TCATCTTCAGGTACACAGCGTTGTGTCTCTGTTGGCTTAAGTAAAGCGG 63
|||||:|||||:|||||:|||||:
DB 623 SerTYRTPArgPheMetGlnCysLeu-----ArgArg 633
|||:|||||:|||||:|||||:
OY 64 -----TGTTCACAGATGATAGACCATGCGACAAACCTGTGACG 105
|||||:|||||:|||||:|||||:
DB 634 PheAlaAspSerGlyAspTrpPheProHisLeuAsnAlaAlaLysTrpThrLeuGly 653
|||:|||||:|||||:|||||:
OY 106 ATTGGCTTTGGATGACCTGTGTGCGTTCTATTGCACAGAAATCAGAG 156
|||:|||||:|||||:|||||:
DB 654 IleAlaTrpAsnAlaThrLeuCysAlaTrpArgLeuSerAspArgSerGlu 670
|||:|||||:|||||:|||||:

RESULT 10
FTR_METTH
ID FTR_METTH STANDARD: PRT; 297 AA.
AC P21348; 027327;
DT 01-MAY-1991 (Rel. 18, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Formylmethanofuran--tetrahydromethanopterin formyltransferase
DE (EC 2.3.1.101).
GN MTH1259 OR FTR.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-61.
RC STRAIN=Delta H;
RA MEDLINE=90094441; PubMed=2403564;
RA DiMarco A.A., Sment K.A., Konisky J., Wolfe R.S.;
RT "The formylmethanofuran:tetrahydromethanopterin formyltransferase
from Methanobacterium thermoautotrophicum delta H. Nucleotide
sequence and functional expression of the cloned gene.";
RL J. Biol. Chem. 265:472-476(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;

RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
RA Alredge T., Baehr-Zadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pochier B., Qiu D.,
RA Spadafora R., Vitare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
delta H: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -I- CATALYTIC ACTIVITY: N-formylmethanofuran + 5,6,7,8-
CC tetrahydromethanopterin = methanofuran + 5-formyl-5,6,7,8-
CC tetrahydromethanopterin.
CC -I- PATHWAY: INVOLVED IN THE FORMATION OF METHANE FROM CO(2).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; J05173; AAA88222.1; -
DR EMBL; AE000892; AAB85748.1; ALT_INIT.
DR PIR; A34912; A34912.
DR HSSP; Q49610; 1FTR.
DR InterPro; IPR002770; FTR.
DR Pfam; PF01913; FTR; 1.
DR Pfam; PF02741; FTR_C; 1.
DR ProDom; PD007702; FTR; 1.
KW Transferase; Acyltransferase; Methanogenesis; Complete proteome.
FT CONFLICT 28 28 MISSING (IN REF. 1).
FT CONFLICT 154 154 E->Q (IN REF. 1).
SQ SEQUENCE 297 AA; 31470 MW; AB9DBDA90952D6 CRC64;

Alignment Scores:
Pred. No.: 14.3 Length: 297
Score: 57.00 Matches: 15
Percent Similarity: 52.50% Conservative: 6
Best Local Similarity: 37.50% Mismatches: 17
Query Match: 17.76% Indels: 2
Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x FTR_METTH (1-297)

OY 122 GTTACTCCAAAGCCATGTCACAGTTTGTGCATGGCTTATCAATTCGAGAAACG 63
|||||:|||||:|||||:|||||:
DB 194 ValThrProPheProGlyGlyValValAlaIleSerGlySerLysValGlySerAsnLysTrp 213
:::|||||:|||||:|||||:|||||:
OY 62 CGCTTCTTAAGTTAGCCACACAGGAAACACAAACCTGTCTTACTTGAAGATGA 3
:::|||||:|||||:|||||:|||||:
DB 214 LysPheLeuAsnAlaSerThr-----AsnGlyLysMetCysValThrLeuLysAspGlu 231
|||:|||||:|||||:|||||:

RESULT 11
MASX_PICAN
ID MASX_PICAN STANDARD: PRT; 555 AA.
AC P21360;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Malate synthase, glyoxysomal (EC 4.1.3.2).
GN MAS.
OS Pichia angusta (Yeast) (Hansenula polymorpha).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4905;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 234-253.
RC STRAIN=MAYA CBS4732;
RX MEDLINE=90273778; PubMed=2349836;
RX Bruinenberg P.G., Blaauw M., Kazemler B., Ab G.;

RA Murphy L., Richards C., Harris D., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 189-1301 FROM N.A.
RC STRAIN-DBY939;
RA Stella C.A., Korch C., Ramos E.H., Mattoon J.R.;
RT "Cloning and sequencing of IEP1, a gene associated with leucine
RT transport.";
RL Yeast 11:460-460(1995).
CC -1- FUNCTION: POTENTIAL REGULATOR OF LEUCINE PERMEASE GENE(S)
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE SAC3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z47805; CAAB767.1; -;
DR EMBL; Z50046; CA90379.1; -;
DR EMBL; U35227; AAA79056.1; -;
DR SGD; S0002566; SAC3.
DR InterPro: IPR005062; SAC3_GANP.
DR Pfam; PF03399; SAC3_GANP; 1.
KW transcription regulation; Nuclear protein.
SQ SEQUENCE 1301 AA; 149568 MW; 0679DB1673DDACEB CRC64;

Alignment Scores:
Pred. No.: 21.5 Length: 1301
Score: 56.00 Matches: 10
Percent Similarity: 57.14% Conservative: 10
Best Local Similarity: 28.57% Mismatches: 15
Query Match: 17.45% Indels: 0
DB: 1 Gaps: 0

US-09-939-293-1_COPY_56_239 (1-184) x SAC3_YEAST (1-1301)
OY 140 ATAGGACCGCACAGGGTACTGTCGAAGCCATCGTCACAGTTTGTGCATGGTCTT 81
||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 623 IIEserGInSerHISThrLeuSerThraSnProLeuLeuThrProGlnValHisGlyAsp 642
OY 80 ATCAATTCGAGAACACCGCTTTTAAAGTTAGCCACACAGCA 36
::: ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 643 IeuSerGluGlnGlyGlnGlnGlnIleuLeuThrValThrAspGly 657

RESULT 14
V084_HSV7J ID V084_HSV7J STANDARD: PRT: 310 AA.
AC P52534.1
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein 084.
GN 084
OS Human herpesvirus (type 7 / strain J1) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-6 AND
CC HSV-7 084 AND HCMV U117.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U43400; AAC54745.1; -;
SQ SEQUENCE 310 AA; 35649 MW; 99BD5B8FBECE7AB5 CRC64;

Alignment Scores:
Pred. No.: 22.8 Length: 310
Score: 55.50 Matches: 15
Percent Similarity: 38.18% Conservative: 6
Best Local Similarity: 27.27% Mismatches: 27
Query Match: 16.97% Indels: 7
DB: 1 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x V084_HSV7J (1-310)
OY 19 AGACAGTGTGTGTGTCTCTGTGTGCTTAACCTTAAAGCGGTGTCTCAGAAATTG 78
::: ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 19 LysSerCysMetCysThrLysAlaSnAlaArgTyrThrCysAsnCysPheSer----- 36
OY 79 ATAAGACCATGGCACAAACTGTGACGATTTGGCTTTGAGATACCTGTGCGGTTCTT 138
||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 37 -----LysThrLeuProPheAsnGluLysAlaIleuLysCysThrIlePro 51
OY 139 ATTGCACGAATCGACAGCCTCATCCCTTAGTAGAAGCATGG 183
||||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 52 GluLysIleAsnSerGluIleAsnIleSerLysSerGluMetIleu 66

RESULT 15
T140_SPTOL ID T140_SPTOL STANDARD: PRT: 449 AA.
AC O49939.1
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peptidyl-prolyl cis-trans isomerase, chloroplast precursor
DE (EC 5.2.1.8) (40 kDa thylakoid lumen peptase) (40 kDa thylakoid lumen
DE rotamase)
GN TLP40 OR TLP40.
OS Spinacia Oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eumariophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC STRAIN=cv. Monatol; TISSUE=green leaf;
RX MEDLINE=98169374; PubMed=9501079;
RA Fulgosi H., Vener A.V., Altschmid L., Herrmann R.G., Andersson B.;
RT "A novel multi-functional chloroplast protein: identification of a 40
RT kDa immunophilin-like protein located in the thylakoid lumen.";
RL EMBO J. 17:1577-1587(1998).
RN [2]
RP SEQUENCE OF 105-134.
RC TISSUE=leaf;
RX MEDLINE=98175931; PubMed=9506969;
RA Kieselbach T., Hagman A., Andersson B., Schroder W.P.;
RT "The thylakoid lumen of chloroplasts. Isolation and
RT characterization.";
RL J. Biol. Chem. 273:6710-6716(1998).
CC -1- FUNCTION: PRIPASES ACCELERATE THE FOLDING OF PROTEINS. HAS A
CC REGULATORY EFFECT ON THYLAKOID PROTEIN PHOSPHORYLATION.
CC -1- CATALYTIC ACTIVITY: peptidylproline (omega-180) - peptidylproline
CC (omega=0).
CC -1- ENZYME REGULATION: DOES NOT BIND CYCLOSPORIN A (CSA).
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST; WITHIN THE THYLAKOID LUMEN AND
CC CAN INTERACT WITH THE INNER SURFACE OF THE STROMA-EXPOSED
CC THYLAKOID REGIONS.
CC -1- SIMILARITY: BELONGS TO THE CYCLOPHILIN-TYPE PRIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Y12071; CAAT2792.1; -
 DR InterPro: IPR002130; CSA_PPIase.
 DR PROSITE: PS00170; CSA_PPIase_1; FALSE_NEG.
 DR PROSITE: PS50072; CSA_PPIase_2; 1.
 KW Cyclosporin; Isomerase; Rotamase; Chloroplast; Transit peptide.
 FT TRANSIT 1 104 CHLOROPLAST.
 FT CHAIN 105 449 PEPTIDYL-PROLYL CIS-TRANS ISOMERASE.
 FT DOMAIN 272 449 PPIASE CYCLOPHILIN-TYPE.
 SO SEQUENCE 449 AA; 49872 MW; C8A40195128F15B1 CRC64;

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-----|
| Pred. No.: | 23.4 | Length: | 449 |
| Score: | 55.50 | Matches: | 14 |
| Percent Similarity: | 44.19% | Conservative: | 5 |
| Best Local Similarity: | 32.56% | Mismatches: | 15 |
| Query Match: | 16.97% | Indels: | 9 |
| DB: | 1 | Gaps: | 1 |

US-09-939-293-1_COPY_56_239 (1-184) x T1A0_SPIOL (1-449)

| | | | |
|----|-----|--|-----|
| QY | 43 | GTGGCTAACTTTAAGAG-----CGGTGTTTCTCAGAA | 75 |
| | ::: | ::: | |
| Db | 32 | IlEProAsnPheArgGlnLysSerArgPheMetHisLeuThrProArgCysPheSerArg | 51 |
| | | | |
| QY | 76 | TTGATTAAGACCATGTCACAAACTGTGACATGTGGCTTTGGAGTAACCCCTGTGCGGTT | 135 |
| | | | |
| Db | 52 | GlnIleAspProLeuAspLysGlnLysLysArgSerPheSerValLysGlnCysAlaIle | 71 |
| | | | |
| QY | 136 | CCTATTGCA | 144 |
| | ::: | | |
| Db | 72 | SerLeuAla | 74 |

Search completed: February 16, 2003, 03:07:35
 Job time : 21.5075 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 16, 2003, 03:05:21 : Search time 26.7761 Seconds
(without alignments)
1321.232 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239
Perfect score: 327
Sequence: 1 actcatctctcagtgatcacg.....ccttagtagtgagacattga 184

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 segs, 96134422 residues
Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p_model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09939293/runat_12022003.170354_8546/app_query.fasta_1.590
-DB=PIR.73 -QFMT=fastan -SUFFIX=rrpr -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MTN=0 -ALICG=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09939293.ecgn_1_1_24.ecrunat_12022003.170354_8546 -NCPG=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LARGOQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR.73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 67 | 20.5 | 202 | 2 T04906 | hypothetical prote |
| 2 | 66.5 | 20.7 | 802 | 2 T24293 | hypothetical prote |
| 3 | 66.5 | 20.7 | 949 | 2 T24294 | hypothetical prote |
| 4 | 65.5 | 20.0 | 2345 | 1 A35378 | acetyl-CoA carboxy |
| 5 | 65.5 | 20.0 | 2346 | 1 A35378 | acetyl-CoA carboxy |
| 6 | 61.5 | 18.8 | 2339 | 2 I38928 | acetyl-CoA carboxy |
| 7 | 60.5 | 18.5 | 322 | 2 H85068 | N7-like protein [i |
| 8 | 60.5 | 18.5 | 322 | 2 H85068 | acetyl-CoA carboxy |
| 9 | 60 | 18.3 | 466 | 2 A29924 | probable K+ transp |
| 10 | 59.5 | 18.5 | 332 | 2 B71314 | probable K+ transp |
| 11 | 58.5 | 18.2 | 351 | 2 B64378 | replication-associ |
| 12 | 58.5 | 18.2 | 351 | 2 B43356 | replication-associ |
| 13 | 58.5 | 17.9 | 902 | 2 S49931 | SYG1 protein - yea |
| 14 | 58.5 | 18.2 | 2233 | 2 T28669 | surface protein 51 |

| | | | | | |
|----|------|------|------|----------|--------------------|
| 15 | 58 | 17.7 | 687 | 2 T04927 | probable serine/th |
| 16 | 58 | 18.1 | 1589 | 2 T13826 | translation initia |
| 17 | 57.5 | 17.6 | 830 | 2 A30359 | P-selectin precurs |
| 18 | 57 | 17.4 | 233 | 2 T01867 | hypothetical prote |
| 19 | 57 | 17.8 | 300 | 2 C69035 | formylmethanofuran |
| 20 | 57 | 17.4 | 555 | 1 SYH0MA | malate synthase (E |
| 21 | 56.5 | 17.3 | 444 | 2 T42979 | hypothetical prote |
| 22 | 56.5 | 17.6 | 816 | 2 C63424 | assimilatory nitri |
| 23 | 56 | 17.1 | 161 | 2 D71892 | probable osmoprote |
| 24 | 56 | 17.1 | 402 | 2 A72077 | hypothetical prote |
| 25 | 56 | 17.1 | 692 | 2 D86547 | hypothetical prote |
| 26 | 56 | 17.1 | 692 | 2 A81593 | hypothetical prote |
| 27 | 56 | 17.4 | 1301 | 2 S51323 | SAC3 protein - yea |
| 28 | 55.5 | 17.3 | 180 | 2 T34851 | probable secreted |
| 29 | 55.5 | 17.0 | 310 | 2 T41985 | hypothetical prote |
| 30 | 55.5 | 17.0 | 449 | 2 T09212 | rotamase Flr40 pre |
| 31 | 55.5 | 17.3 | 504 | 2 T21377 | hypothetical prote |
| 32 | 55.5 | 17.0 | 522 | 2 C75448 | glycerol-3-phospha |
| 33 | 55.5 | 17.0 | 1496 | 2 T45808 | helicase-like prot |
| 34 | 55.5 | 17.0 | 3133 | 2 S52093 | hemocytin - silkw |
| 35 | 55 | 16.8 | 162 | 2 E84455 | probable RING zinc |
| 36 | 55 | 16.8 | 374 | 2 D86965 | probable glycosylt |
| 37 | 55 | 16.8 | 445 | 2 T31581 | hypothetical prote |
| 38 | 55 | 16.8 | 449 | 2 B85069 | hypothetical prote |
| 39 | 55 | 17.1 | 503 | 2 E98159 | choline sulfatase |
| 40 | 55 | 17.1 | 503 | 2 AD3128 | choline sulfatase |
| 41 | 55 | 17.1 | 636 | 2 F69027 | cleavage and polya |
| 42 | 55 | 17.1 | 727 | 2 T37748 | sepin interacting |
| 43 | 55 | 16.8 | 1798 | 2 S53869 | lamnin beta-2 cha |
| 44 | 55 | 16.8 | 1888 | 2 T14273 | zinc finger protei |
| 45 | 54.5 | 16.7 | 212 | 2 S51798 | gamma-kafirin prec |

ALIGNMENTS

RESULT 1

T04906

hypothetical protein T10114.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 02-Sep-2000

C:Accession: T04906

R:By:van, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.

submitted to the Protein Sequence Database, April 1998

A:Reference number: Z15389

A:Accession: T04906

A:Molecule type: DNA

A:Residues: 1-202 <BEV>

A:Cross-references: EMBL:AL021712

A:Experimental source: cultivar Columbia; BAC clone T10114

C:Genetics:

A:Map position: 4

A>Note: T10114.80

C:Superfamily: RING finger homology

F:150-197/Domain: RING finger homology <RRN>

Alignment Scores:

Pred. No.: 2.25 Length: 202
Score: 67.00 Matches: 19
Percent Similarity: 44.23% Conservative: 4
Best Local Similarity: 36.54% Mismatches: 24
Query Match: 20.49% Indels: 5
DB: 2 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x T04906 (1-202)

QY 11 TCAGGTACGACAGTGTGTGCTTCCTGCTGCTACCTTAAGAAGCGGTGTTCT 70

DB 148 TTTGTAASASPSerValCys-CysValCysMetGlyArgLysGlyAlaAlaPhe1 167

QY 71 CAGATTGATTAGACCATGGACAAACCTGTCAGTCATGGCTTTCGAGTAAC----- 122

DB 167 eetrocysglnshtrhphncysArgValCysSerArgGluLeuTrpLeuAsnArgGlyse 187

```

Oy 123 -----CCTGTGCGGTTCTTATGACAGAAATC 152
      |||::|||
Db 187 rCysProLeuCYsAsnArGProIleIleGluLe 198

RESULT 2
T24293
hypotheetical protein T01D3.3a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24293
R:Steward, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19870
A:Accession: T24293
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-802 <Wtl>
A:Cross-references: EMBL:Z81110; P1DN:CA803259.1; GSPDB:GN00023; CESP:T01D3.3a
A:Experimental source: clone T01D3
C:Genetics:
A:Gene: CESP:T01D3.3a
A:Map position: 5
A:Introns: 74/1; 121/1; 200/2; 493/1; 673/3; 772/2

Alignment Scores:
Pred. No.: 2.45 Length: 802
Score: 66.50 Matches: 19
Percent Similarity: 46.67% Conservative: 9
Best local Similarity: 31.67% Mismatches: 27
Query Match: 20.72% Indels: 5
Db: 2 Gaps: 2

US-09-939-293-1_COPY_56_239 (1-184) x T24293 (1-802)
Oy 182 AATGCTTACACTAGAGGAATGAGGCTCTGATTTCTGCAATAGAACCCGACACAGG 123
      |||::|||
Db 225 AsnProGlyMetIleGlySerPheGlyAsnThrAsnIleGlnTyrProThrSerProArg 244

Oy 122 GTTACTCCAAAGCCAAATGTCACAGTTTGTGCCATGCTTATCAATTCAGAAACAC 63
      |||||::|||
Db 245 AlaThrProSerProValPheThrThrThrArgGlyLeuThrThrSerGlnGln---- 263

Oy 62 CGCTTCTTAAGTTAGCCACACAGGAACACAAACCTGTCTGTAC---CTGAAGAT 6
      |||||::|||
Db 264 -----LysLeuProThrLeuIleProGluGlnHisCysGluHisProLeuLysAsn 280

RESULT 3
T24294
hypotheetical protein T01D3.3b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24294
R:Steward, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19870
A:Accession: T24294
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-949 <Wtl>
A:Cross-references: EMBL:Z81110; P1DN:CA803260.1; GSPDB:GN00023; CESP:T01D3.3b
A:Experimental source: clone T01D3
C:Genetics:
A:Gene: CESP:T01D3.3b
A:Map position: 5
A:Introns: 74/1; 126/1; 161/2; 221/1; 268/1; 347/2; 640/1; 820/3; 919/2

Alignment Scores:
Pred. No.: 2.43 Length: 949
Score: 66.50 Matches: 19
Percent Similarity: 46.67% Conservative: 9
Best local Similarity: 31.67% Mismatches: 27
Query Match: 20.72% Indels: 5
Db: 2 Gaps: 2

US-09-939-293-1_COPY_56_239 (1-184) x T24294 (1-949)
Oy 182 AATGCTTACACTAGAGGAATGAGGCTCTGATTTCTGCAATAGAACCCGACACAGG 123
      |||::|||
Db 372 AsnProGlyMetIleGlySerPheGlyAsnThrAsnIleGlnTyrProThrSerProArg 391

Oy 122 GTTACTCCAAAGCCAAATGTCACAGTTTGTGCCATGCTTATCAATTCAGAAACAC 63
      |||||::|||
Db 392 AlaThrProSerProValPheThrThrThrArgGlyLeuThrThrSerGlnGln---- 410

Oy 62 CGCTTCTTAAGTTAGCCACACAGGAACACAAACCTGTCTGTAC---CTGAAGAT 6
      |||||::|||
Db 411 -----LysLeuProThrLeuIleProGluGlnHisCysGluHisProLeuLysAsn 427

RESULT 4
A35578
acetyl-CoA carboxylase (EC 6.4.1.2) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Feb-2002
C:Accession: A35578; A37119; I59145; I70069; I70070; I55305
R:Lopez-Casillas, F.; Bai, D.H.; Luo, X.; Kong, I.S.; Hermodson, M.A.; Kim, K.H.
Proc. Natl. Acad. Sci. U.S.A. 85, 5784-5788, 1988
A:Title: Structure of the coding sequence and primary amino acid sequence of acetyl-c
A:Reference number: A35578; MUID:88320328; PMID:2901088
A:Accession: A35578
A:Molecule type: mRNA
A:Residues: 1-2345 <LOP>
A:Cross-references: GB:J03808; NID:g202644; P1DN:AAA40653.1; PID:g202645
R:Kong, I.S.; Lopez-Casillas, F.; Kim, K.H.
J. Biol. Chem. 265, 13695-13701, 1990
A:Title: Acetyl-CoA carboxylase mRNA species with or without inhibitory coding sequen
A:Reference number: A37119; MUID:90337981; PMID:1974251
A:Accession: A37119
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1167-1200 <KON>
A:Cross-references: GB:M55315
A:Experimental source: clone lambdaDHN121
R:Luo, X.
Proc. Natl. Acad. Sci. U.S.A. 86, 4042-4046, 1989
A:Title: Structural features of the acetyl-CoA carboxylase gene: Mechanism for the ge
A:Reference number: I59145; MUID:89264558; PMID:2566999
A:Accession: I59145
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-74 <RES>
A:Cross-references: GB:M26731; NID:g202641; P1DN:AAA40652.1; PID:g554406
A:Experimental source: hepatic
R:Lopez-Casillas, F.; Kim, K.
J. Biol. Chem. 264, 7176-7184, 1989
A:Title: Heterogeneity at the 5' end of rat acetyl-coenzyme A carboxylase mRNA: Lipog
A:Reference number: I55305; MUID:89214151; PMID:2565337
A:Accession: I70070
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-33 <RE3>
A:Cross-references: GB:M26197; NID:g202651; P1DN:AAA40656.1; PID:g554409
A:Accession: I55305
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-33 <RE4>
A:Cross-references: GB:M26195; NID:g202647; P1DN:AAA40654.1; PID:g554407
A:Experimental source: hepatic
C:Comment: This enzyme catalyzes the carboxylation of acetyl CoA to malonyl CoA and i
C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biot
C:Keywords: biotin binding; lipase
F:119-619/Domain: biotin carboxylase homology <BCH>

```

F:746-818/Domain: lipoyl/biotin-binding homology <LBP>
F:785/Binding site: biotin (lys) (covalent) #status predicted

Alignment Scores:

Pred. No.: 3.15 Length: 2345
Score: 65.50 Matches: 15
Percent Similarity: 51.02% Mismatches: 10
Best Local Similarity: 30.61% Conservative: 13
Query Match: 20.03% Indels: 11
DB: 1 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x A35578 (1-2345)

QY 67 TTCTCGAATTGATAGACGACCAAACTGTCAGATTGGC----- 111
|||||.....
DB 1959 PheSerGIuIleMetInProtrPalaGIuThrValValaValGIuYrGAlaArgLeuGIy 1978
QY 112 -----TTTGAGTAACCCCTGTGCGGTCCTATTTCACAGAAATCA 153
|||.....
DB 1979 GIuIleProValaGIuValValaValaGIuThrArgTrhValGIuLeuSerValProAla 1998
QY 154 GAGCCTCATTCCTTAGTAGTAAGCA 180
:::|||||.....
DB 1999 AspProAlaAsnLeuAspSerGIuaLa 2007

RESULT 5

138928
acetyl-CoA carboxylase (EC 6.4.1.2), hepatic - human
N:Alternate names: acetyl-Coenzyme A carboxylase
C:Species: Homo sapiens (man)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 11-Jan-2002
C:Accession: J38928
R:Abu-Elheida, L.; Jayakumar, A.; Baldini, A.; Chirala, S.S.; Wakil, S.U.
Proc. Natl. Acad. Sci. U.S.A. 92, 4011-4015, 1995
A>Title: Human acetyl-CoA carboxylase: characterization, molecular cloning, and evidence
A:Reference number: 138928; MUID:95249602; PMID:7732023
A:Accession: J38928
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2346 <RES>
A:Cross-references: EMBL:U19822; NID:g849082; PIDN:AAC50139.1; PID:g849083
A:Experimental source: HepG2 cells
C:Genetics:
A:Gene: ACC
A:Map position: 17q12
C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C:Keywords: biotin binding; ligase; liver
F:120-620/Domain: biotin carboxylase homology <BC>
F:747-819/Domain: lipoyl/biotin-binding homology <LBP>
F:786/Binding site: biotin (lys) (covalent) #status predicted

Alignment Scores:

Pred. No.: 3.15 Length: 2346
Score: 65.50 Matches: 15
Percent Similarity: 51.02% Mismatches: 10
Best Local Similarity: 30.61% Conservative: 13
Query Match: 20.03% Indels: 11
DB: 2 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x I38928 (1-2346)

QY 67 TTCTCGAATTGATAGACGACCAAACTGTCAGATTGGC----- 111
|||||.....
DB 1960 PheSerGIuIleMetInProtrPalaGIuThrValValaValGIuYrGAlaArgLeuGIy 1979
QY 112 -----TTTGAGTAACCCCTGTGCGGTCCTATTTCACAGAAATCA 153
|||.....
DB 1980 GIuIleProValaGIuValValaValaGIuThrArgTrhValGIuLeuSerValProAla 1999
QY 154 GAGCCTCATTCCTTAGTAGTAAGCA 180
:::|||||.....
DB 2000 AspProAlaAsnLeuAspSerGIuaLa 2008

RESULT 6
S41121
acetyl-CoA carboxylase (EC 6.4.1.2) - human

C:Species: Homo sapiens (man)
C>Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 11-Jan-2002
C:Accession: S41121
R:Ha, J.; Daniel, S.; Kong, I.S.; Park, C.K.; Tae, H.J.; Kim, K.H.
Eur. J. Biochem. 219, 297-306, 1994
A>Title: Cloning of human acetyl-CoA carboxylase cDNA.
A:Reference number: S41121; MUID:94139704; PMID:7905825
A:Accession: S41121
A:Molecule type: mRNA
A:Residues: 1-2339 <HND>
A:Cross-references: EMBL:568968; NID:9452315; PIDN:CAA8770.1; PID:9452316
C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C:Keywords: biotin binding; ligase
F:120-620/Domain: biotin carboxylase homology <BC>
F:747-819/Domain: lipoyl/biotin-binding homology <LBP>
F:786/Binding site: biotin (lys) (covalent) #status predicted

Alignment Scores:

Pred. No.: 10.7 Length: 2339
Score: 61.50 Matches: 16
Percent Similarity: 42.59% Mismatches: 7
Best Local Similarity: 28.63% Conservative: 20
Query Match: 18.81% Indels: 11
DB: 2 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x S41121 (1-2339)

QY 52 TTTAAGACGGGCTTTCACAGATTGATAGACCAAACTGTCAGATTGGC 111
|||.....
DB 1947 PheAspHisGIuSerPheIuysGIuIleMetAlaProtrPalaGIuThrValaValaThrGIy 1966
QY 112 -----TTTGAGTAACCCCTGTGCGGTCCT 138
|||.....
DB 1967 ArgAlaArgLeuGIyGIuIleProValaGIuValaValaGIuThrArgTrhValGIu 1986
QY 139 ATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGTAAGCA 180
:::|||||.....
DB 1987 ValAlaValaProAlaAspProAlaAsnLeuAspSerGIuaLa 2000

RESULT 7

H85068
N7-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: H85068
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp.
Nature 402, 769-777, 1999
A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85068
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-322 <STO>
A:Cross-references: GB:NC_001268; NID:g7267308; PIDN:CAB81090.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4G05480
A:Map position: 4

Alignment Scores:

Pred. No.: 15.9 Length: 322
Score: 60.50 Matches: 14
Percent Similarity: 50.00% Mismatches: 4
Best Local Similarity: 38.89% Conservative: 15
Query Match: 18.50% Indels: 3
DB: 2 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x H85068 (1-322)

QY 19 AGACAGCTTTTGTGTCCTGTCCTGCTTGAAGCGGCTT-----TTC 69
|||||.....

C:Species: wheat dwarf virus
 A:Variety: French isolate
 C:Date: 16-Feb-1995 #sequence_revision 06-Dec-1996 #text_change 21-Aug-1998
 C:Accession: S49387
 R:Bendahmane, M.; Schalk, H.J.; Gronenborn, B.
 Submitted to the EMBL Data Library, October 1994
 A:Description: Identification and characterization of wheat dwarf virus (WDV) from France
 A:Reference number: S49385
 A:Accession: S49387
 A:Molecule type: DNA
 A:Residues: 1-351 <BN>
 A:Cross-references: EMBL:X82104
 A:Experimental source: French isolate
 A:Note: ORF C1-1 and ORF C1-2 are joined by removal of an intron in the region of overlap
 R:Schalk, H.J.; Matzeit, V.; Schiller, B.; Schell, J.; Gronenborn, B.
 EMO J. 8, 359-364, 1989
 A:Title: Wheat dwarf virus, a geminivirus of graminaceous plants needs splicing for replication
 A:Reference number: S71838; MID:89251559; PMID:2721484
 A:Contents: annotation; Intron position
 C:Genetics:
 A:Gene: C1
 A:Introns: 210/3
 C:Superfamily: tomato golden mosaic virus AL1 protein
 C:Keywords: DNA replication

Alignment Scores:
 Pred. No.: 29.2 Length: 351
 Score: 58.50 Matches: 14
 Percent Similarity: 44.44% Conservative: 6
 Best Local Similarity: 31.11% Mismatches: 16
 Query Match: 18.22% Indels: 9
 Gaps: 1
 DB: 2

US-09-939-293-1_COPY_56_239 (1-184) x S49387 (1-351)

QY 134 ACCGACACAGGGTTACTCCAAAGCCATGTCACA----- 99
 ||||| :||| ||||| |||
 Db 173 ThAlaArgHisLeuPheProAspProValAlaThrTyrThrProGluPheProThrglu 192

QY 98 ---GTTTGGCCATGCTTATCAATTCGAGAAACCGCTTTAAAGTTAGCCACA 42
 :|||:||||| ||| ||| ||||| |||
 Db 193 SerLeuIleCysHisGluThrIleGluSerTrpLysAsnIuHisLeuTyrSerGluSer 212

QY 41 ACAGAACACACAAA 27
 ||| |||||
 Db 213 ProGlyArgHisLys 217

RESULT 12
 B24356
 Replication-associated protein - wheat dwarf virus (Swedish isolate)
 N:Alternate names: ORF 30156/ORF 17292 composite protein
 C:Species: wheat dwarf virus
 A:Variety: Swedish isolate
 C:Date: 09-Sep-1987 #sequence_revision 06-Dec-1996 #text_change 21-Aug-1998
 C:Accession: B24356
 R:MacDowell, S.W.; MacDonald, H.; Hamilton, W.D.O.; Coutts, R.H.A.; Buck, K.W.
 EMO J. 4, 2173-2180, 1985
 A:Title: The nucleotide sequence of cloned wheat dwarf virus DNA.
 A:Reference number: A91012
 A:Accession: B24356
 A:Molecule type: DNA
 A:Residues: 1-351 <MAC>
 A:Cross-references: GB:X02869
 A:Experimental source: Swedish isolate
 A:Note: ORF 30156 and ORF 17292 are joined by removal of an intron in the region of overlap
 R:Schalk, H.J.; Matzeit, V.; Schiller, B.; Schell, J.; Gronenborn, B.
 EMO J. 8, 359-364, 1989
 A:Title: Wheat dwarf virus, a geminivirus of graminaceous plants needs splicing for replication
 A:Reference number: S71838; MID:89251559; PMID:2721484
 A:Contents: annotation; Intron position
 C:Genetics:

A:Introns: 210/3
 C:Superfamily: tomato golden mosaic virus AL1 protein
 C:Keywords: DNA replication

Alignment Scores:
 Pred. No.: 29.2 Length: 351
 Score: 58.50 Matches: 14
 Percent Similarity: 44.44% Conservative: 6
 Best Local Similarity: 31.11% Mismatches: 16
 Query Match: 18.22% Indels: 9
 Gaps: 1
 DB: 2

US-09-939-293-1_COPY_56_239 (1-184) x B24356 (1-351)

QY 134 ACCGACACAGGGTTACTCCAAAGCCATGTCACA----- 99
 ||||| :||| ||||| |||
 Db 173 ThAlaArgHisLeuPheProAspProValAlaThrTyrThrProGluPheProThrglu 192

QY 98 ---GTTTGGCCATGCTTATCAATTCGAGAAACCGCTTTAAAGTTAGCCACA 42
 :|||:||||| ||| ||| ||||| |||
 Db 193 SerLeuIleCysHisGluThrIleGluSerTrpLysAsnIuHisLeuTyrSerGluSer 212

QY 41 ACAGAACACACAAA 27
 ||| |||||
 Db 213 ProGlyArgHisLys 217

RESULT 13
 S49331
 SYG1 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein Y19905.01c; protein Y1L047c
 C:Species: Saccharomyces cerevisiae
 C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 06-Feb-1998
 C:Accession: S49331; S48245
 R:Odell, C.; Bowman, S.
 Submitted to the EMBL Data Library, December 1994
 A:Reference number: S49331
 A:Accession: S49331
 A:Molecule type: DNA
 A:Residues: 1-902 <ODP>
 A:Cross-references: GB:Z47047; EMBL:Z46861; MID:9603997; PID:9763299; MIPS:Y1L047c
 R:Spain, B.H.; Koo, D.; Ramakrishnan, M.; Dzubozor, B.; Colicelli, J.
 Submitted to the EMBL Data Library, September 1994
 A:Description: Truncations of a novel gene can suppress the lethality of a C-alpha su
 A:Reference number: S48245
 A:Accession: S48245
 A:Molecule type: DNA
 A:Residues: 1-176, 178-902 <SPA>
 A:Cross-references: EMBL:U14726
 C:Genetics:
 A:Gene: SGD:SYG1
 A:Cross-references: SGD:S0001309; MIPS:Y1L047c
 A:Map position: 9L
 C:Keywords: transmembrane protein
 F:401-417/Domain: transmembrane #status predicted <TM1>
 F:444-460/Domain: transmembrane #status predicted <TM2>
 F:498-514/Domain: transmembrane #status predicted <TM3>
 F:527-543/Domain: transmembrane #status predicted <TM4>
 F:555-571/Domain: transmembrane #status predicted <TM5>
 F:581-597/Domain: transmembrane #status predicted <TM6>
 F:766-782/Domain: transmembrane #status predicted <TM7>

Alignment Scores:
 Pred. No.: 27.8 Length: 902
 Score: 58.50 Matches: 16
 Percent Similarity: 45.61% Conservative: 10
 Best Local Similarity: 28.07% Mismatches: 16
 Query Match: 17.89% Indels: 15
 Gaps: 2
 DB: 2

US-09-939-293-1_COPY_56_239 (1-184) x S49331 (1-902)

QY 4 TCAATCTTCAGTACAGACAGTGTGTGTCCTGTGTGTCCTAATTAAAGACGG 63
 |||:|||||:||||| ||||| |||||
 :|||:|||||:||||| ||||| |||||

Db 623 SerTyrTPArgPheMetGlnCysLeu-----ArgArg 633
 QY 64 -----TGTTCACGAATGATAGACCATGGCAAAACTGTGACG 105
 Db 634 PheAlaAspSerGlyAspTrpPheProHisLeuAsnAlaAlaLysTyrThrLeuGly 653
 QY 106 APTGGCTTTGGATGACCCGTGTGGCGTTCCTATTGCACAGAAATCAGAG 156
 Db 654 IleAlaTyrAsnAlaThrLeuCysAlaTyrArgLeuSerAspArgSerGlu 670

RESULT 14

T28669

surface protein 51C - Paramecium tetraurelia

C:Species: Paramecium tetraurelia

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999

C:Accession: T28669

R.Nielsen, E., You, Y., Forney, J.

J. Mol. Biol. 222, 835-841, 1991

A:Title: Cysteine residue periodicity is a conserved structural feature of variable surf

A:Reference number: 220504; MUID:92106337; PMID:1762150

A:Accession: T28669

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2233 <NTE>

A:Cross-References: EMBL:M65164; NID:q159974; PID:q159975; PIDN:AAA61740.1

C:Genetics:

A:Genetic code: SGCS

Alignment Scores:

Pred. No.: 26.6 Length: 2233
 Score: 58.50 Matches: 18
 Percent Similarity: 52.38% Conservative: 4
 Best Local Similarity: 42.86% Mismatches: 15
 Query Match: 18.22% Indels: 5
 DB: 2 Gaps: 3

US-09-939-293-1_COPY_56_239 (1-184) x T28669 (1-2233)

QY 158 GGCCTGATTTCTGTCGCAATGAGACCGACACAGGGT---ACTCCAAAGCCATCTGC 102
 Db 702 GlyAsnProLeuCysValValGlyThrAlaAsnAsnValCysAlaProLeuPro----- 719
 QY 101 ACAGTTTGTGCCATGGCTCT--ATCAATTCGAGAACACCGCTTCTTAAAGTTAGCC 45
 Db 720 ---AlaThrCysSerGlyLeuGlyLysAsnSerCysLysThrAsnIleLeuLysValAla 738
 QY 44 ACAACA 39
 Db 739 AsnThr 740

RESULT 15

T04927

probable serine/threonine-specific protein kinase (EC 2.7.1.-) T9A21.100 - Arabidopsis t

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999

C:Accession: T04927

R.Beyan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15390

A:Accession: T04927

A:Molecule type: DNA

A:Residues: 1-687 <BEV>

A:Cross-References: EMBL:AL021713

A:Experimental source: cultivar Columbia; BAC clone T9A21

C:Genetics:

A:Map position: 4

A:Introns: 223/1; 299/3

A>Note: T9A21.100

C:Keywords: phosphotransferase; protein kinase

Alignment Scores:

Pred. No.: 32.9 Length: 687
 Score: 58.00 Matches: 15

Percent Similarity: 47.06% Conservative: 9
 Best Local Similarity: 29.41% Mismatches: 19
 Query Match: 17.74% Indels: 8
 DB: 2 Gaps: 2

US-09-939-293-1_COPY_56_239 (1-184) x T04927 (1-687)

QY 1 ACTGATTTCTGAGGTACAGACAG--TGTGTGTGTCTGTGTGGCTAACTTTAAG 57
 Db 240 ThrAsnTyrPheLysTyrLysGlnThrCysLysProThrProTyrValGlnAsnAsp 259
 QY 58 AAGCGGTGTTTCAGATTGATAAGACCATGGCAAAA----- 96
 Db 260 ArgAlaCysProSerAlaTyrSerTyrProPheSerGlyAsnAsnSerThrPheThrCys 279
 QY 97 ACTGTGACGATGGCTTGGAGTAACCCGTGT 129
 Db 280 ThrAsnSerThrAspTyrValIleThrPheCys 290

Search completed: February 16, 2003, 03:11:59
 Job time : 30.7761 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 00:27:34 : Search time 88.5672 Seconds
(without alignments)
1058.130 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 184

Sequence: 1 attcattcttcagtgacag.....ccttagtagtagagcattga 184

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 424239 seqs, 254651826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PC7_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 184 | 100.0 | 1358 | 10 | US-09-939-293-1 |
| 2 | 183 | 99.5 | 911 | 10 | US-09-925-297-132 |
| 3 | 182.4 | 99.1 | 549 | 9 | US-09-796-692-8602 |
| 4 | 122 | 66.3 | 714 | 10 | US-09-798-116-3 |
| 5 | 122 | 66.3 | 1356 | 10 | US-09-798-116-1 |
| 6 | 30.6 | 16.6 | 32203 | 10 | US-09-764-869-1849 |
| 7 | 29.8 | 16.2 | 403 | 10 | US-09-983-965-5387 |
| 8 | 29.6 | 16.1 | 36159 | 12 | US-10-135-687-3 |
| 9 | 29.4 | 16.0 | 1059 | 9 | US-09-911-345-1 |
| 10 | 28.8 | 15.7 | 779 | 9 | US-09-822-846-14 |
| 11 | 28.8 | 15.7 | 817 | 10 | US-09-941-437-1 |
| 12 | 28.8 | 15.7 | 2838 | 10 | US-09-960-253-158 |
| 13 | 28.8 | 15.7 | 6892 | 10 | US-09-764-869-1720 |
| 14 | 28.2 | 15.3 | 369 | 10 | US-09-783-590-4715 |
| 15 | 28.2 | 15.3 | 8459 | 10 | US-09-817-913-8 |
| 16 | 28.2 | 15.3 | 8459 | 10 | US-09-817-913-8 |
| 17 | 28 | 15.2 | 371 | 10 | US-09-815-343-856 |
| 18 | 27.8 | 15.1 | 546 | 10 | US-09-974-300-5879 |
| 19 | 27.8 | 15.1 | 1617 | 10 | US-09-815-242-7028 |

| | | | | | | | |
|---|----|------|------|--------|----|---------------------|-------------------|
| c | 20 | 27.8 | 15.1 | 7573 | 10 | US-09-880-107-2195 | Sequence 2195, Ap |
| c | 21 | 27 | 14.7 | 4015 | 9 | US-10-071-766-113 | Sequence 113, App |
| c | 22 | 26.8 | 14.6 | 309 | 10 | US-09-728-446-238 | Sequence 238, App |
| c | 23 | 26.8 | 14.6 | 1059 | 9 | US-09-911-345-3 | Sequence 3, App11 |
| c | 24 | 26.8 | 14.6 | 17993 | 10 | US-09-768-781-5 | Sequence 5, App11 |
| c | 25 | 26.6 | 14.5 | 27681 | 10 | US-09-764-869-1997 | Sequence 1997, Ap |
| c | 26 | 26.6 | 14.5 | 27681 | 10 | US-09-764-869-1998 | Sequence 1998, Ap |
| c | 27 | 26.4 | 14.3 | 483 | 10 | US-09-833-381-605 | Sequence 605, App |
| c | 28 | 26.4 | 14.3 | 7478 | 9 | US-10-104-966-15 | Sequence 15, App1 |
| c | 29 | 26.4 | 14.3 | 7478 | 10 | US-09-929-955-15 | Sequence 15, App1 |
| c | 30 | 26.2 | 14.2 | 498 | 10 | US-09-864-761-14348 | Sequence 14348, A |
| c | 31 | 26.2 | 14.2 | 29449 | 9 | US-09-989-442-161 | Sequence 161, App |
| c | 32 | 26 | 14.1 | 534 | 10 | US-09-798-116-5 | Sequence 5, App11 |
| c | 33 | 26 | 14.1 | 804 | 9 | US-09-938-824-3398 | Sequence 3398, Ap |
| c | 34 | 26 | 14.1 | 1052 | 9 | US-09-969-844-8 | Sequence 8, App1 |
| c | 35 | 26 | 14.1 | 1999 | 10 | US-09-864-761-3802 | Sequence 3802, Ap |
| c | 36 | 26 | 14.1 | 7043 | 10 | US-09-070-927A-288 | Sequence 288, App |
| c | 37 | 26 | 14.1 | 13526 | 10 | US-09-764-887-645 | Sequence 645, App |
| c | 38 | 26 | 14.1 | 14337 | 10 | US-09-764-887-644 | Sequence 644, App |
| c | 39 | 26 | 14.1 | 53542 | 10 | US-09-801-574-61 | Sequence 61, App1 |
| c | 40 | 25.8 | 14.0 | 342 | 10 | US-09-974-300-3776 | Sequence 3776, Ap |
| c | 41 | 25.8 | 14.0 | 451 | 10 | US-09-833-381-287 | Sequence 287, App |
| c | 42 | 25.8 | 14.0 | 3720 | 9 | US-09-822-846-180 | Sequence 180, App |
| c | 43 | 25.8 | 14.0 | 9274 | 10 | US-09-885-535-3 | Sequence 3, App11 |
| c | 44 | 25.8 | 14.0 | 167343 | 10 | US-09-962-436-281 | Sequence 281, App |
| c | 45 | 25.8 | 14.0 | 167343 | 10 | US-09-964-824A-273 | Sequence 273, App |

ALIGNMENTS

RESULT 1
US-09-939-293-1
; Sequence 1, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alnemat, Emed S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; TITLE OF INVENTION: AND METHODS OF USING THE SAME
; FIDE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1358
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)...(739)
US-09-939-293-1

Query Match 100.0%; Score 184; DB 10; Length 1358;

Best Local Similarity 100.0%; Pred. No. 3.6e-51;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 1 | ACTTCATCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAG | 60 |
| DB | 56 | ACTTCATCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAG | 115 |
| QY | 61 | CGGTGTTCTTCAGAACTTGAAGACCAATGCGCAAAACGTATACCATTTGGCTTGAAGTA | 120 |
| DB | 116 | CGGTGTTCTTCAGAACTTGAAGACCAATGCGCAAAACGTATACCATTTGGCTTGAAGTA | 175 |
| QY | 121 | ACCCGTGTGCGGCTTCCTATTCACAGAAATAGAGCCCTTCCCTTAGTAGTAAGCA | 180 |
| DB | 176 | ACCCGTGTGCGGCTTCCTATTCACAGAAATAGAGCCCTTCCCTTAGTAGTAAGCA | 235 |
| QY | 181 | TTGA 184 | |
| DB | 236 | TTGA 239 | |

```

: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8602
: LENGTH: 549
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-796-692-8602

Query Match          99.1%: Score 182.4; DB 9; Length 549;
Best Local Similarity 99.5%: Pred. No. 8.7e-51;
Matches 183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ACTCATCTCTCAGGTACAGACAGTGTGTGTGTCTCTCTGTGTGCTTAACCTTAAGAG 60
Db       51  ACTCATCTCTCAGGTACAGACAGTGTGTGTGTCTCTGTGTGCGCTTAACCTTAAGAG 110
QY      61  CGGTGTCTCTAGAACTTATTAAGACCATGCGACAAACTGTGACGATTTGCTTGAGAGTA 120
Db      111  CGGTGTCTCTAGAACTTATTAAGACCATGCGATTAACCTGTGACGATTTGCTTGAGAGTA 170
QY     121  ACCCGTGTGGGGTTCCTATTGACACAGAAATCAGAGCCTCATTCCTAGTAAGTAAGCA 180
Db     171  ACCCGTGTGGGGTTCCTATTGACACAGAAATCAGAGCCTCATTCCTAGTAAGTAAGCA 230
QY      181  TTGA 184
Db     231  TTGA 234

RESULT 4
US-09-798-116-3
: Sequence 3, Application US/09798116
: Patent No. US20020110851A1
: GENERAL INFORMATION:
: APPLICANT: Verhagen, Anne Marie
: APPLICANT: Ekerl, Paul
: APPLICANT: Vaux, David
: TITLE OF INVENTION: No. US20020110851A1 Polypeptides, Modulatory Agents Therefor
: FILE REFERENCE: 10338-004US
: CURRENT FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: US/09/798,116
: PRIOR FILING DATE: 2000-03-02
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 3
: LENGTH: 714
: TYPE: DNA
: ORGANISM: Mus musculus
US-09-798-116-3

66.3%: Score 122; DB 10; Length 714;

```

; SOFTWARE: PatentIn Ver. 2.0

PRIOR FILING DATE: 2000-12-28

: NUMBER OF SEQ ID NOS: 4
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO: 3
 : LENGTH: 36159
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: (1)...(36159)
 : OTHER INFORMATION: n - A,T,C or G
 US-10-135-687-3

Query Match 16.1%; Score 29.6; DB 12; Length 36159;
 Best Local Similarity 51.5%; Pred. No. 10;
 Matches 68; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 28 TTGCTGCTCTGTTGGCTTAACCTTTAAGACGGGTCTTCTCAGATTGATAAGACCA 87
 Db 12881 TTCTTAATTTTCAAGAGCTCTAACAATGTGTAATCTTAATCTTATGATGACACCA 12940
 QY 88 TGGCAACAACCTGACGAGTGGCTTGGAGTAACCTGTGCGGTTCTATTGACACAG 147
 Db 12941 TTAATAATCTTGTGATAGATTTGTGTAGATTAAGTACCTACACATTATTAAGAACTT 13000
 QY 148 AATTCAGACCT 159
 Db 13001 AGAACAGTGCAAT 13012

RESULT 9
 US-09-911-345-1
 : Sequence 1, Application US/09911345
 : Patent No. US20020160449A1
 : GENERAL INFORMATION:
 : APPLICANT: Holloway, James L.
 : APPLICANT: Iok, Si
 : TITLE OF INVENTION: Human Vomeronasal Receptor
 : FILE REFERENCE: 00-45
 : CURRENT APPLICATION NUMBER: US/09/911,345
 : CURRENT FILING DATE: 2001-07-23
 : NUMBER OF SEQ ID NOS: 3
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO: 1
 : LENGTH: 1059
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (1)...(1059)
 US-09-911-345-1

Query Match 16.0%; Score 29.4; DB 9; Length 1059;
 Best Local Similarity 52.9%; Pred. No. 3.2;
 Matches 63; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 24 GTGTTGTGTTCTCTGTTGGCTAACTTTAAGACGGGTGTTCTCAGATTGATAG 83
 Db 829 GTCTCGTGAGACCTCTTTTGTTCATTCAGTCACTTAATTTGACATTGGACCA 888
 QY 84 ACCATGGACAACACTGAGAGATTGCTTGGAGTAACCTGTGTGGGTTCTTATG 142
 Db 889 ACTGTAGTGCACCAACGACGACAGTAGTACCAACTCTGTGTGGTGGCTTCATG 947

RESULT 10
 US-09-822-846-14
 : Sequence 14, Application US/09822846
 : Publication No. US20030027139A1
 : GENERAL INFORMATION:
 : APPLICANT: Jacobs, Kenneth
 : APPLICANT: McCoy, John M.
 : APPLICANT: Lavallee, Edward R.
 : APPLICANT: Collins-Racie, Lisa A.

: APPLICANT: Evans, Cheryl
 : APPLICANT: Merberg, David
 : APPLICANT: Treacy, Maurice
 : APPLICANT: Agostino, Michael J.
 : APPLICANT: Steinger II, Robert J.
 : APPLICANT: Bowman, Michael R.
 : APPLICANT: Spaulding, Vikki
 : APPLICANT: Wong, Gordon G.
 : APPLICANT: Clark, Hilary
 : APPLICANT: Fechtel, Kim
 : APPLICANT: Howes, Steven H.
 : APPLICANT: Resnick, Richard J.
 : APPLICANT: Gulukota, Kamalakara
 : APPLICANT: Graham, James R.
 : APPLICANT: Genetics Institute, Inc.
 : TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
 : FILE REFERENCE: GIN 6400
 : CURRENT APPLICATION NUMBER: US/09/822,846
 : CURRENT FILING DATE: 2001-03-29
 : PRIOR APPLICATION NUMBER: 60/195,605
 : PRIOR FILING DATE: 2000-04-06
 : NUMBER OF SEQ ID NOS: 629
 : SOFTWARE: Patent Ver. 2.0
 : SEQ ID NO: 14
 : LENGTH: 779
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-822-846-14

Query Match 15.7%; Score 28.8; DB 9; Length 779;
 Best Local Similarity 58.0%; Pred. No. 4.5;
 Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 14 GGTACAGACAGTGTGTTGGTCTCTGTTGGCTAATTTAAGACGGGTGTTCTCAG 73
 Db 410 GTTCAGAGATTGTTGTTCTTTCTAGACTTGGGATCTGCAGACGCCAATTGCCCTAA 465
 QY 74 AATTGATTAAGACATGACCAAACTGT 101
 Db 470 AATTCTGAGAACAGTGACCAAGATTAT 497

RESULT 11
 US-09-941-437-1
 : Sequence 1, Application US/09941437
 : Patent No. US20020106356A1
 : GENERAL INFORMATION:
 : APPLICANT: Jacobs, Kenneth
 : APPLICANT: McCoy, John M.
 : APPLICANT: Racie, Lisa A.
 : APPLICANT: Lavallee, Edward R.
 : APPLICANT: Merberg, David
 : APPLICANT: Treacy, Maurice
 : APPLICANT: Spaulding, Vikki
 : APPLICANT: Agostino, Michael

: TITLE OF INVENTION: SECRETED PROTEINS
 : NUMBER OF SEQUENCES: 2
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Genetics Institute, Inc.
 : STREET: 87 Cambridgepark Drive
 : CITY: Cambridge
 : STATE: MA
 : COUNTRY: U.S.A.
 : ZIP: 02140
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/941,437
 : FILING DATE: 28-Aug-2001
 : CLASSIFICATION: <Unknown>

```
;
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/885,469
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: P-41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 817 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-941-437-1

Query Match          15.7%; Score 28.8; DB 10; Length 817;
Best Local Similarity 58.0%; Pred. No. 4.6;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 14 GGTACAGACAGTGTGTGTCTGTGCTGTGCGCTAATTAGAAAGCGGTGTTCTCAG 73
DB 414 GTTGACAGTGTGTGTCTGTGCTGTGCGCTAATTAGAAAGCGGTGTTCTCAG 473

QY 74 AATTGATAGACCATGGCACAATACTGT 101
DB 474 AATTCTGAGAACAGTGCCACAAAGATTAT 501

RESULT 12
US-09-960-253-158
; Sequence 158, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.556
; CURRENT APPLICATION NUMBER: US/09/960,253
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 158
; LENGTH: 2838
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-158

Query Match          15.7%; Score 28.8; DB 10; Length 2838;
Best Local Similarity 58.0%; Pred. No. 7.4;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 14 GGACAGACAGTGTGTGTCTGTGCTGTGCGCTAATTAGAAAGCGGTGTTCTCAG 73
DB 2341 GTTGACAGTGTGTGTGTCTGTGCTGTGCGCTAATTAGAAAGCGGTGTTCTCAG 2400

QY 74 AATTGATAGACCATGGCACAATACTGT 101
DB 2401 AATTCTGAGAACAGTGCCACAAAGATTAT 2428

RESULT 13
US-09-764-869-1720
; Sequence 1720, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
;
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1720
; LENGTH: 6892
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1720

Query Match          15.7%; Score 28.8; DB 10; Length 6892;
Best Local Similarity 58.0%; Pred. No. 10;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 14 GGTACAGACAGTGTGTGTCTGTGCTGTGCGCTAATTAGAAAGCGGTGTTCTCAG 73
DB 6386 GTTGACAGTGTGTGTGTCTGTGCTGTGCGCTAATTAGAAAGCGGTGTTCTCAG 6445

QY 74 AATTGATAGACCATGGCACAATACTGT 101
DB 6446 AATTCTGAGAACAGTGCCACAAAGATTAT 6473

RESULT 14
US-09-783-590-4715
; Sequence 4715, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16,201
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4715
; LENGTH: 369
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (24)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (36)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (60)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (63)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (66)
```

```
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (73)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (125)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (173)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (176)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (195)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (199)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (217)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (227)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (230)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (238)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (251)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (268)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (269)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (300)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (304)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (308)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (310)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (323)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (353)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (358)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-4715
```

```
Query Match 15.3%; Score 28.2; DB 10; Length 369;
Best Local Similarity 55.8%; Pred. No. 5.4;
Matches 48; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
```

```
QY 99 TGTGAGATTGGCTTGGAGTAACCTGTGCGGTTCCATTGACAGAAATCAGAGCC 158
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 136 TTGGAAGATTGCTGCTAAGGCTCCCTGTTATCTNCTNTTGGCCAGAACTGTGTN 195
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 159 TCATTCCCTTACTAGTGAAGCATTTGA 184
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Db 196 TCCNACGTTAAATGTAAGTCTGA 221

```
RESULT 15
US-09-817-913-8/c
; Sequence 8, Application US/09817913
; Patent No. US20020061860A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zuomei
; APPLICANT: Bonfils, Claire
; APPLICANT: Besterman, Jeffrey
; TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms
; FILE REFERENCE: 106101.145
; CURRENT APPLICATION NUMBER: US/09/817,913
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,157
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 8459
; TYPE: DNA
; ORGANISM: Human
US-09-817-913-8
```

```
Query Match 15.3%; Score 28.2; DB 10; Length 8459;
Best Local Similarity 64.6%; Pred. No. 17;
Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```
QY 35 TTCCTGTTGTGGCTAAGCTTTAAGAGCGGTGTTCTCAGAAATGATTAAGACCAGGCACA 94
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 4869 TTCCTGCTGTTGTTTCTTAAGGCAATGTTCTCAGAACTGGAAGAAACCCAGGCACA 4810
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 95 AACT 99
    | | | | |
DB 4809 AAAAT 4805
```

Search completed: February 16, 2003, 03:05:10
Job time : 109.567 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 12:12:34 : Search time 63.8507 Seconds
(without alignments)
883.757 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 184

Sequence: 1 actcattcttcaggtacag.....ccttagtagtggaagatttga 184

Scoring table: IDENTITY_NUC

Searched: 443362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 184 | 100.0 | 720 | 3 | US-09-479-309-1 |
| 2 | 169.6 | 92.2 | 720 | 3 | US-09-479-309-3 |
| 3 | 169.6 | 92.2 | 720 | 3 | US-09-479-309-4 |
| 4 | 169.6 | 92.2 | 720 | 3 | US-09-479-309-5 |
| 5 | 155.2 | 84.3 | 720 | 3 | US-09-479-309-6 |
| 6 | 155.2 | 84.3 | 720 | 3 | US-09-479-309-7 |
| 7 | 153.2 | 83.3 | 720 | 3 | US-09-479-309-8 |
| 8 | 30 | 16.3 | 2383 | 4 | US-09-221-017B-629 |
| 9 | 29.6 | 16.1 | 36159 | 4 | US-09-749-588-3 |
| 10 | 28.8 | 15.7 | 817 | 4 | US-08-885-469-1 |
| 11 | 28.8 | 15.7 | 1405 | 4 | US-09-625-918-1 |
| 12 | 28.8 | 15.7 | 1405 | 4 | US-09-149-476-183 |
| 13 | 27.8 | 15.1 | 873 | 4 | US-09-134-001C-2081 |
| 14 | 27.8 | 15.1 | 7573 | 1 | US-08-287-959-2 |
| 15 | 27.8 | 15.1 | 368 | 4 | US-08-651-155B-211 |
| 16 | 26.6 | 14.5 | 943 | 1 | US-07-807-043B-12 |
| 17 | 26.6 | 14.5 | 943 | 1 | US-08-299-849B-12 |
| 18 | 26.6 | 14.5 | 943 | 2 | US-08-142-368A-12 |
| 19 | 26.6 | 14.5 | 943 | 3 | US-08-967-727-12 |
| 20 | 26.6 | 14.5 | 943 | 4 | US-08-037-230D-12 |
| 21 | 26.6 | 14.1 | 1052 | 4 | US-09-592-891A-8 |
| 22 | 25.6 | 13.9 | 945 | 4 | US-09-056-105-7 |
| 23 | 25.6 | 13.9 | 1019 | 4 | US-09-056-105-13 |
| 24 | 25.6 | 13.9 | 1375 | 2 | US-08-993-738A-2 |
| 25 | 25.6 | 13.9 | 1375 | 4 | US-08-713-354C-2 |
| 26 | 25.6 | 13.9 | 1375 | 4 | US-09-241-268-2 |
| 27 | 25.6 | 13.9 | 1375 | 4 | US-09-495-562-2 |

| | | | | | | |
|----|------|------|------|---|---------------------|-------------------|
| 28 | 25.6 | 13.9 | 1640 | 1 | US-07-807-043B-11 | Sequence 11, Appl |
| 29 | 25.6 | 13.9 | 1640 | 2 | US-08-299-849B-11 | Sequence 11, Appl |
| 30 | 25.6 | 13.9 | 1640 | 3 | US-08-142-368A-11 | Sequence 11, Appl |
| 31 | 25.6 | 13.9 | 1640 | 3 | US-08-967-727-11 | Sequence 11, Appl |
| 32 | 25.6 | 13.9 | 1640 | 4 | US-08-037-230D-11 | Sequence 11, Appl |
| 33 | 25.6 | 13.9 | 2709 | 4 | US-09-134-001C-2320 | Sequence 2320, Ap |
| 34 | 25.6 | 13.9 | 4204 | 2 | US-08-928-615-1 | Sequence 1, Appl |
| 35 | 25.6 | 13.9 | 4204 | 4 | US-09-056-105-6 | Sequence 6, Appl |
| 36 | 25.6 | 13.9 | 4204 | 4 | US-09-166-448-1 | Sequence 1, Appl |
| 37 | 25.6 | 13.9 | 4204 | 4 | US-09-348-933-1 | Sequence 1, Appl |
| 38 | 25.6 | 13.9 | 4204 | 4 | US-09-697-884-1 | Sequence 1, Appl |
| 39 | 25.4 | 13.8 | 1997 | 2 | US-08-969-630-1 | Sequence 1, Appl |
| 40 | 25.2 | 13.7 | 1968 | 4 | US-08-961-527-119 | Sequence 119, App |
| 41 | 25.2 | 13.7 | 2443 | 2 | US-08-685-625A-1 | Sequence 1, Appl |
| 42 | 25.2 | 13.7 | 3273 | 1 | US-08-148-122A-2 | Sequence 2, Appl |
| 43 | 25.2 | 13.7 | 5503 | 4 | US-09-453-702B-195 | Sequence 195, App |
| 44 | 25 | 13.6 | 429 | 3 | US-08-945-809A-56 | Sequence 56, Appl |
| 45 | 25 | 13.6 | 645 | 4 | US-09-328-111-9 | Sequence 9, Appl |

ALIGNMENTS

| | | |
|--|---|-------------------------------------|
| RESULT 1 | US-09-479-309-1 | Sequence 1, Application US/09479309 |
| Patent No. 6110691 | | |
| GENERAL INFORMATION: | | |
| APPLICANT: Wang, Xiaodong | | |
| APPLICANT: Du, Chunying | | |
| TITLE OF INVENTION: Activators of Caspases | | |
| FILE REFERENCE: UTS0630 | | |
| CURRENT APPLICATION NUMBER: US/09/479,309 | | |
| CURRENT FILING DATE: 2000-01-06 | | |
| NUMBER OF SEQ ID NOS: 8 | | |
| SOFTWARE: PatentIn Ver. 2.1 | | |
| SEQ ID NO 1 | | |
| LENGTH: 720 | | |
| TYPE: DNA | | |
| ORGANISM: human | | |
| FEATURE: | | |
| NAME/KEY: CDS | | |
| LOCATION: (1)..(717) | | |
| US-09-479-309-1 | | |
| Query Match | 100.0%; | Score 184; DB 3; Length 720; |
| Best Local Similarity | 100.0%; | Pred. No. 5.6e-57; |
| Matches 184; Conservative | 0; | Mismatches 0; Indels 0; Gaps 0; |
| QY 1 | ACTTCATTCCTTCAGTACAGACAGTGTGTGTGTCCTGTCGCTACTTAAAGG 60 | |
| DB 37 | ACTTCATTCCTTCAGTACAGACAGTGTGTGTGTCCTGTCGCTACTTAAAGG 96 | |
| QY 61 | CGGTGTTTCAGAAATGATAGACATGCGACAAACTGTGACGATGGCTTGGAGTA 120 | |
| DB 97 | CGGTGTTTCAGAAATGATAGACATGCGACAAACTGTGACGATGGCTTGGAGTA 156 | |
| QY 121 | ACCTGTTGCGGTTCCATTGACAGAAATCAGAGCCTTACCTTACTAGTAAGCA 180 | |
| DB 157 | ACCTGTTGCGGTTCCATTGACAGAAATCAGAGCCTTACCTTACTAGTAAGCA 216 | |
| QY 181 | TTGA 184 | |
| DB 217 | TTGA 220 | |
| RESULT 2 | US-09-479-309-3 | |
| Sequence 3, Application US/09479309 | | |
| Patent No. 6110691 | | |
| GENERAL INFORMATION: | | |
| APPLICANT: Wang, Xiaodong | | |
| APPLICANT: Du, Chunying | | |

```

; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTS0630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-3

```

```

Query Match          92.2%; Score 169.6; DB 3; Length 720;
Best Local Similarity 95.1%; Pred. No. 8.8e-52;
Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

QY 1 ACTTCATCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAACTTTAAGAG 60
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 37 ACTTCATCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAACTTTAAGAG 96
QY 61 CGGTGTTTCTCAGAAATGATTAAGACCATGCGACAAAACGTGTGACGATTGGAGTA 120
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 97 CGGTGTTTCTCAGAAATGATTAAGACCATGCGACAAAACGTGTGACGATTGGAGTA 156
QY 121 ACCCTGTGTGCGGTCCTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAAGA 180
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 157 ACCCTGTGTGCGGTCCTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAAGA 216
QY 181 TTGA 184
    ||||
DB 217 TTGA 220

```

```

RESULT 3
US-09-479-309-4
; Sequence 4, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTS0630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-4

```

```

Query Match          92.2%; Score 169.6; DB 3; Length 720;
Best Local Similarity 95.1%; Pred. No. 8.8e-52;
Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

QY 1 ACTTCATCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAACTTTAAGAG 60
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 37 ACTTCATCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAACTTTAAGAG 96
QY 61 CGGTGTTTCTCAGAAATGATTAAGACCATGCGACAAAACGTGTGACGATTGGAGTA 120
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 97 CGGTGTTTCTCAGAAATGATTAAGACCATGCGACAAAACGTGTGACGATTGGAGTA 156
QY 121 ACCCTGTGTGCGGTCCTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAAGA 180
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 157 ACCCTGTGTGCGGTCCTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAAGA 216

```

```

QY 181 TTGA 184
    ||||
DB 217 TTGA 220

```

```

RESULT 4
US-09-479-309-5
; Sequence 5, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTS0630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-5

```

```

Query Match          92.2%; Score 169.6; DB 3; Length 720;
Best Local Similarity 95.1%; Pred. No. 8.8e-52;
Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

QY 1 ACTTCATCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAACTTTAAGAG 60
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 37 ACTTCATCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAACTTTAAGAG 96
QY 61 CGGTGTTTCTCAGAAATGATTAAGACCATGCGACAAAACGTGTGACGATTGGAGTA 120
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 97 CGGTGTTTCTCAGAAATGATTAAGACCATGCGACAAAACGTGTGACGATTGGAGTA 156
QY 121 ACCCTGTGTGCGGTCCTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAAGA 180
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 157 ACCCTGTGTGCGGTCCTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTGTAGTAGAAGA 216
QY 181 TTGA 184
    ||||
DB 217 TTGA 220

```

```

RESULT 5
US-09-479-309-6
; Sequence 6, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTS0630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 6
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-6

```

```

Query Match          84.3%; Score 155.2; DB 3; Length 720;
Best Local Similarity 90.2%; Pred. No. 1.4e-46;

```


TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 629:
SEQUENCE CHARACTERISTICS:
LENGTH: 2383 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...2383
US-09-221-017B-629

Query Match 16.3%; Score 30; DB 4; Length 2383;
Best Local Similarity 54.5%; Pred. No. 0.35;
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 24 GTGTTGTCCTCTGTTGGCTTAAGACGGGTGTTCTCAGATTGATAG 83
DB 2058 GTTATTCTGTTCTTTAGCTGATGAGATGAGACTGTTCAAGATTCATAG 2117
QY 84 ACCATGCGACAAACTGACGATGCTTGGAGTACCTGTCGCG 133
DB 2118 CAATTGCCGACGACCTGAACCAATTTCTGGGAAAAATGATGCTTGGCG 2167

RESULT 9
US-09-749-588-3
Sequence 3, Application US/09749588
Patent No. 6423521
GENERAL INFORMATION:
APPLICANT: CHANDRAMOULISWARAN, Ishwar et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINSE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CU001068
CURRENT APPLICATION NUMBER: US/09/749,588
CURRENT FILING DATE: 2000-12-28
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 36159
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(36159)
OTHER INFORMATION: n = A,T,C or G
US-09-749-588-3

Query Match 16.1%; Score 29.6; DB 4; Length 36159;
Best Local Similarity 51.5%; Pred. No. 1.6;
Matches 68; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 28 TTGTCGTCCTCTGTTGGCTTAAGACGGGTGTTCTCAGATTGATAGACCA 87
DB 12881 TTCTTAATTCTCAGAGGTCAATCAATGTAATAAAGATGTATCTTGAATTGATGACATCA 12940
QY 88 TGGCACAACACTGTGACATGGCTTGGAGTAACCTGTGCGGTCCTCATTCACAG 147
DB 12941 TTAATATCTTGTGATGAGATTGTGTAGATTAGTACATCAACATTATTAAGAACTT 13000
QY 148 AAATCAGAGCT 159
DB 13001 AGACAGTGCAAT 13012

RESULT 10
US-08-885-469-1
Sequence 1, Application US/08885469
Patent No. 6280739
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Racie, Lisa A.
APPLICANT: Lavallee, Edward R.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: AGOSTINO, Michael
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/885,469
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 817 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-885-469-1

Query Match 15.7%; Score 28.8; DB 4; Length 817;
Best Local Similarity 58.0%; Pred. No. 0.59;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 14 GTACAGACAGTGTGTTGTCCTCTGTTGGCTTAAGACGGGTGTTCTCAG 73
DB 414 GTTCAGATGATGTTGTTCTTTCTAGACTTGGATCTGCAAGAGCCAAATGCCCTAA 473
QY 74 AATTGATTAAGACATGCGACAAAACGT 101
DB 474 AATTCTGAGAACAGTGCACAAAGATTAT 501

RESULT 11
US-09-625-918-1
Sequence 1, Application US/09625918
Patent No. 6451318
GENERAL INFORMATION:
APPLICANT: JACOBS, Kenneth
APPLICANT: MCCOY, John M.
APPLICANT: RACIE, Lisa A.
APPLICANT: LAVALLIE, Edward R.
APPLICANT: MERBERG, David
APPLICANT: TREACY, Maurice
APPLICANT: SPAULDING, Vikki
APPLICANT: AGOSTINO, Michael
TITLE OF INVENTION: SECRETED PROTEINS

```

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSFE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/625,918
FILING DATE: 26-Jul-2000
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 817 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
GS-09-625-918-1

```

```

Query Match      15.7%: Score 28.8; DB 4; Length 817;
Best Local Similarity 58.0%: Pred. No. 0.59;
Matches 51: Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 14 GGTACAGACAGCTGTTGTGTGTCTCTGTGTGGCTAACTTTAAGACGCGGTCTTCTCAG 73
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 414 GTTCGAGTAGTGGTTGGTTGCTTTCTTTCTAGACCTTGGGATCTGCAGAAAGCCAAFTTGCCTAA 473
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 74 AATTGATTAAGACCATGGCACAANAAGTCT 101
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 474 AATTCTGAGAAACAGTGCACAAGATTAT 501
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-149-476-183
; Sequence 183, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: 1998-03-06
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600

```

| | | |
|---------|---------------------|------------|
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,615 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,597 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,502 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,633 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,583 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,617 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,618 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,503 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,584 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,500 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,587 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,613 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,582 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,586 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,612 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,611 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,600 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/043,580 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,568 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,314 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,569 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,311 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,671 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,674 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,669 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,312 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,313 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,672 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,315 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/048,974 |
| EARLIER | FILING DATE: | 1997-06-06 |
| EARLIER | APPLICATION NUMBER: | 60/056,886 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,877 |
| EARLIER | FILING DATE: | 1997-08-22 |

| | |
|-----------------------------|------------|
| EARLIER APPLICATION NUMBER: | 60/056,889 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,893 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,630 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,878 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,662 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,872 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,882 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,637 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,903 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,888 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,879 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,880 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,894 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,911 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,636 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,874 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,910 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,864 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/056,631 |
| EARLIER FILING DATE: | 1997-08-22 |
| EARLIER APPLICATION NUMBER: | 60/047,550 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,594 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,566 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,588 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,585 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,586 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,550 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,594 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,589 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,593 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/047,614 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/043,578 |
| EARLIER FILING DATE: | 1997-04-11 |
| EARLIER APPLICATION NUMBER: | 60/043,576 |
| EARLIER FILING DATE: | 1997-04-11 |
| EARLIER APPLICATION NUMBER: | 60/047,501 |
| EARLIER FILING DATE: | 1997-05-23 |
| EARLIER APPLICATION NUMBER: | 60/043,670 |
| EARLIER FILING DATE: | 1997-04-11 |
| EARLIER APPLICATION NUMBER: | 60/056,632 |

```

1 EARLIER FILING DATE: 1997-08-22
2 EARLIER APPLICATION NUMBER: 60/056,664
3 EARLIER FILING DATE: 1997-08-22
4 EARLIER APPLICATION NUMBER: 60/056,876
5 EARLIER FILING DATE: 1997-08-22
6 EARLIER APPLICATION NUMBER: 60/056,883
7 EARLIER FILING DATE: 1997-08-22
8 EARLIER APPLICATION NUMBER: 60/056,909
9 EARLIER FILING DATE: 1997-08-22
10 EARLIER APPLICATION NUMBER: 60/056,875
11 EARLIER FILING DATE: 1997-08-22
12 EARLIER APPLICATION NUMBER: 60/056,862
13 EARLIER FILING DATE: 1997-08-22
14 EARLIER APPLICATION NUMBER: 60/056,887
15 EARLIER FILING DATE: 1997-08-22
16 EARLIER APPLICATION NUMBER: 60/056,908
17 EARLIER FILING DATE: 1997-08-22
18 EARLIER APPLICATION NUMBER: 60/048,964
19 EARLIER FILING DATE: 1997-06-06
20 EARLIER APPLICATION NUMBER: 60/057,650
21 EARLIER FILING DATE: 1997-09-05
22 EARLIER APPLICATION NUMBER: 60/056,884
23 EARLIER FILING DATE: 1997-08-22
24 EARLIER APPLICATION NUMBER: 60/057,665
25 EARLIER FILING DATE: 1997-09-05
26 EARLIER APPLICATION NUMBER: 60/049,610
27 EARLIER FILING DATE: 1997-06-13
28 EARLIER APPLICATION NUMBER: 60/061,060
29 EARLIER FILING DATE: 1997-10-02

```

| | | | | |
|-----------------------|--------|-----------------|--------|----------------|
| Query Match | 15.7%; | Score 28.8; | DB 4; | Length 1405; |
| Best Local Similarity | 58.0%; | Pred. No. 0.75; | | |
| Matches | 51; | Conservative | 0; | Mismatches 37; |
| | | | Indels | 0; |
| | | | Gaps | 0; |

OY 14 GGCACGACACTGTTTGGTGCTTCCTGTTGGCTACTTTAAGACGGGTTCAC 73
 Db 881 GTTCGACGATTTGGTTGTTCTTTTCTAGACTTGGATCTGCAGAGGCCATTGGCTAA 940
 OY 74 AATTGATTAAGACCATGGCACAACCTGT 101
 Db 941 AATTCTGAGAACAGTGCACCAAGTTTAT 968

```

RESULT 13
US-09-134-001C-2081
: Sequence 2081, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucelte-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GNC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 2081
: LENGTH: 873
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2081

```

| | | | | | | | | | | | |
|-----|-------------|--|-------|-------|----------|------------|-----|--------|----|------|----|
| | Query Match | Similarity | 59.5% | Pred. | No. 1.4: | Mismatches | 32: | Indels | 0: | Gaps | 0: |
| bdb | 89 | GGCACAAACTGTGAGCATGGCTTTTGAGATACCCTGTGCGTTCCTATTGCACAGA | 148 | | | | | | | | |
| | 419 | GACCAATTAATAATGCCGATGTCTTCTTGAGATTAAGCCGGCGGTGACAATATGATGAAGCATTAC | 478 | | | | | | | | |

THIS PAGE BLANK (USPTO)

[illegible]

XX Claim 1; SEQ ID 11276; 71pp + CD-ROM; English.
PS
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly(A⁺ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 325 BP; 71 A; 75 C; 86 G; 92 T; 1 other;
Query Match 100.0%; Score 184; DB 21; Length 325;
Best Local Similarity 100.0%; Pred. No. 8.2e-53;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTTCATCTTCAGGTACAGACAGCTGTTGTGTCTCTGTTGGCTAACTTAAAG 60
DB 88 ACTTCATCTTCAGGTACAGACAGCTGTTGTGTCTCTGTTGGCTAACTTAAAG 147
QY 61 CGGTGTTCTCAGAAATGATTAAGACATGACAAACTGAGAGATTGGCTTGAAGTA 120
DB 148 CGGTGTTCTCAGAAATGATTAAGACATGACAAACTGAGAGATTGGCTTGAAGTA 207
QY 121 ACCCTGTGGGGTCTCTATTGACAGAAATCAGAGCCTCATTCCTTAGTAGTGAACA 180
DB 208 ACCCTGTGGGGTCTCTATTGACAGAAATCAGAGCCTCATTCCTTAGTAGTGAACA 267
QY 181 TTGA 184
DB 268 TTGA 271

RESULT 2
ID AAA94860 standard; cDNA; 720 BP.
XX
AC AAA94860;
XX
DT 23-FEB-2001 (first entry)
XX
DE Human caspase activator Smac coding sequence.
XX
KW Human; caspase activator; Smac; apoptosis; cancer; autoimmune disease;
KW neurodegenerative disease; mitochondria; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 1..720
FT /*tag= a
FT /product= "Smac"
XX
XX US6110691-A.
XX
PD 29-AUG-2000.
XX
PF 06-JAN-2000; 2000US-0479309.
XX
PR 06-JAN-2000; 2000US-0479309.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
FI Wang X, Du C;
XX

DR WPI; 2000-586350/55.
DR P-PSDB; AAB26210.
XX
PT Novel caspase regulatory polypeptides useful for screening binding
PT agents specific for the polypeptides which are useful for diagnosis and
PT also for treating apoptosis associated diseases -
XX
PS Claim 5; column 21-22; 16pp; English.
XX
CC The present sequence is the coding sequence of the human Smac protein. It
CC was isolated by purifying the protein and searching a HeLa cell cDNA
CC library for sequences which is bound to probes based upon it. Smac is a
CC mitochondrial protein which is released into the cytosol during
CC apoptosis, and acts as a caspase-3 activator. The coding sequence and its
CC protein can be used to modulate the expression and function of caspases
CC and their activators, and also can be used as drug targets and regulators
CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
CC and neurodegenerative diseases.
XX
SQ Sequence 720 BP; 206 A; 147 C; 195 G; 172 T; 0 other;
Query Match 100.0%; Score 184; DB 21; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.1e-52;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTTCATCTTCAGGTACAGACAGCTGTTGTGTCTCTGTTGGCTAACTTAAAG 60
DB 37 ACTTCATCTTCAGGTACAGACAGCTGTTGTGTCTCTGTTGGCTAACTTAAAG 96
QY 61 CGGTGTTCTCAGAAATGATTAAGACATGACAAACTGAGAGATTGGCTTGAAGTA 120
DB 97 CGGTGTTCTCAGAAATGATTAAGACATGACAAACTGAGAGATTGGCTTGAAGTA 156
QY 121 ACCCTGTGGGGTCTCTATTGACAGAAATCAGAGCCTCATTCCTTAGTAGTGAACA 180
DB 157 ACCCTGTGGGGTCTCTATTGACAGAAATCAGAGCCTCATTCCTTAGTAGTGAACA 216
QY 181 TTGA 184
DB 217 TTGA 220

RESULT 3
ID ABR15451 standard; DNA; 1358 BP.
XX
AC ABR15451;
XX
DT 18-JUN-2002 (first entry)
XX
DE DNA encoding inhibitor of apoptosis (IAP) protein Smac.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; tumour; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 20..739
FT /*tag= a
FT /product= "Smac protein"
XX
XX WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX

PI Alnemri ES:
 XX WPI: 2002-304115/34.
 DR P-PSDB; AAU78447.
 XX
 PT Novel Smac peptides and polynucleotides encoding the peptides, useful
 PT for stimulating apoptosis in neoplastic or tumour cell which
 PT overexpresses inhibitor of caspase, and for identifying apoptosis
 PT modulating compounds -
 XX
 PS Claim 1: Page 73-74; 78pp: English.
 XX
 CC The invention relates to an isolated Smac peptide or polypeptide (I)
 CC and an isolated nucleic acid (II) encoding (I). Also described is a
 CC method of identifying a compound that inhibits apoptosis, comprising:
 CC (a) separately contacting several cell populations expressing a
 CC cytosolic Smac (a Smac isoform that begins with MKSPFY sequence,
 CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
 CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
 CC domain) with a compound to be tested for apoptotic inhibiting activity;
 CC (b) incubating the cell populations with a direct stimulus of the cell
 CC death pathway; and (c) measuring the specific apoptotic activity of the
 CC cell populations, where inhibition of the specific apoptotic activity is
 CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
 CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
 CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
 CC tumour cell which overexpresses an inhibitor of caspase, where the
 CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
 CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
 CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
 CC mediated apoptosis which involves contacting a cell transformed or
 CC transfected with a vector expressing (I) with a candidate inhibitor or
 CC candidate enhancer; and detecting cell viability, where an increase in
 CC cell viability indicates the presence of an inhibitor and a decrease in
 CC cell viability indicates the presence of an enhancer. Optionally, the
 CC method involves detecting the presence of large and small caspase
 CC subunits after contacting cell transformed with the vector expressing
 CC (I), with the candidate compound. A decrease in processing indicates the
 CC presence of an inhibitor and an increase in the processing indicates the
 CC presence of an enhancer. Preferably, the large and small subunits of
 CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
 CC identifying a compound that inhibits Smac binding to Smac-binding
 CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
 CC or a full-length IAP). (II) is useful in gene therapy techniques. The
 CC present sequence represents the coding sequence of Smac protein.
 CC
 XX
 SQ Sequence 1358 BP; 341 A; 314 C; 358 G; 345 T; 0 other;
 Query Match 100.0%; Score 184; DB 24; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 1.4e-52;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC AAC98904:
 XX
 DT 09-MAR-2001 (first entry)
 XX
 DE Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:132.
 XX
 KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW detection; diagnosis; identification; cytostatic; neuroprotective;
 KW neurotropic; immunomodulatory; relaxant; contraceptive; gynaecological;
 KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200055320-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-0505989.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI: 2000-579444/54.
 DR P-PSDB; AAB54139.
 XX
 PT New nucleic acid that is a pancreatic cancer antigen for preventing,
 PT treating, or ameliorating a medical condition, particular pancreatic
 PT cancer, or for use in assays for diagnosing a pathological condition -
 XX
 PS Claim 1: Page 591; 1379pp: English.
 XX
 CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,
 CC neuroprotective, neurotropic, immunomodulatory, relaxant, contraceptive,
 CC gynaecological, cardiant and antiinflammatory activities, and can be used
 CC in gene therapy. The polynucleotide and proteins can be used for
 CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in chromosome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect and target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.
 CC
 XX
 SQ Sequence 911 BP; 239 A; 202 C; 253 G; 211 T; 6 other;
 Query Match 99.5%; Score 183; DB 21; Length 911;
 Best Local Similarity 100.0%; Pred. No. 2.6e-52;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
 AAC98904
 ID AAC98904 standard; cDNA; 911 BP.
 XX

DB 62 GGTGTTCTCAGATTGATAAGACCATGGCAAAACCTGTGACGATTGGCTTTGGAGTAA 121
 124 GTGTCTCTCAGATTGATAAGACCATGGCAAAACCTGTGACGATTGGCTTTGGAGTAA 183

OY 122 CCTGTGTCGGTTCCTATTGACAGAAATCAGAGCCCTCATTCCTTAGTAGTAAGCAT 181
 |||||
 DB 184 CCTGTGTCGGTTCCTATTGACAGAAATCAGAGCCCTCATTCCTTAGTAGTAAGCAT 243

OY 182 TGA 184
 ||||
 DB 244 TGA 246

RESULT 5

AAA94247
 ID AAA94247 standard; DNA; 720 BP.

AC AAA94247;

DT 23-FEB-2001 (first entry)

DE Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 3.

XX Human; caspase activator; Smac; hybridising sequence; apoptosis; cancer;
 KW autoimmune disease; neurodegenerative disease; mitochondria; ss.
 XX

OS Homo sapiens.

PN US6110691-A.

PD 29-AUG-2000.

PF 06-JAN-2000; 2000US-0479309.

PR 06-JAN-2000; 2000US-0479309.

PA (TEXA) UNIV TEXAS SYSTEM.

PI Wang X, Du C;

DR WPI; 2000-586350/55.

PT Novel caspase regulatory polypeptides useful for screening binding
 PT agents specific for the polypeptides which are useful for diagnosis and
 also for treating apoptosis associated diseases -

PS Disclosure; column 23-26; 16pp; English.

CC The present sequence is a sequence which was shown to hybridise to the
 CC coding sequence of the human Smac protein. The Smac coding sequence
 CC was isolated by purifying the protein and searching a HeLa cell cDNA
 CC library for sequences which bound to probes based upon it. Smac is a
 CC mitochondrial protein which is released into the cytosol during
 CC apoptosis, and acts as a caspase-3 activator. The coding sequence and
 CC protein can be used to modulate the expression and function of caspases
 CC and their activators, and also can be used as drug targets and regulators
 CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
 CC and neurodegenerative diseases.
 CC
 XX

SO Sequence 720 BP; 207 A; 145 C; 190 G; 178 T; 0 other;

Query Match 92.2%; Score 169.6; DB 21; Length 720;
 Best Local Similarity 95.1%; Pred. No. 8.8e-48;
 Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 ACTTCATTCTTCAAGTACAGACAGTGTGTCGTCCTGTTGCTAGCTACTTAAGAAG 60
 |||||
 DB 37 ACTTCATTCTTCAAGTACAGACAGTGTGTCGTCCTGTTGCTAGCTACTTAAGAAG 96

OY 61 CGGTGTTTTCAGAAATGATTAAGACCATGGCAAAATGTCAGCTTGGCTTGAAGA 120
 |||||
 DB 97 CGGTGTTTTCAGAAATGATTAAGACCATGGCAAAATGTCAGCTTGGCTTGAAGA 156

OY 121 ACCCTGTGTGCGGTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTAGTAGTAAGA 180
 |||||
 DB 157 ACCCTGTGTGCGGTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTAGTAGTAAGA 216

OY 181 TTGA 184
 ||||
 DB 217 TTGA 220

RESULT 6
 AAA94248
 ID AAA94248 standard; DNA; 720 BP.

AC AAA94248;

DT 23-FEB-2001 (first entry)

DE Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 4.

XX Human; caspase activator; Smac; hybridising sequence; apoptosis; cancer;
 KW autoimmune disease; neurodegenerative disease; mitochondria; ss.
 XX

OS Homo sapiens.

PN US6110691-A.

PD 29-AUG-2000.

PF 06-JAN-2000; 2000US-0479309.

PR 06-JAN-2000; 2000US-0479309.

PA (TEXA) UNIV TEXAS SYSTEM.

PI Wang X, Du C;

DR WPI; 2000-586350/55.

PT Novel caspase regulatory polypeptides useful for screening binding
 PT agents specific for the polypeptides which are useful for diagnosis and
 also for treating apoptosis associated diseases -

PS Disclosure; column 25-26; 16pp; English.

CC The present sequence is a sequence which was shown to hybridise to the
 CC coding sequence of the human Smac protein. The Smac coding sequence
 CC was isolated by purifying the protein and searching a HeLa cell cDNA
 CC library for sequences which bound to probes based upon it. Smac is a
 CC mitochondrial protein which is released into the cytosol during
 CC apoptosis, and acts as a caspase-3 activator. The coding sequence and
 CC protein can be used to modulate the expression and function of caspases
 CC and their activators, and also can be used as drug targets and regulators
 CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
 CC and neurodegenerative diseases.
 CC
 XX

SO Sequence 720 BP; 202 A; 154 C; 192 G; 172 T; 0 other;

Query Match 92.2%; Score 169.6; DB 21; Length 720;
 Best Local Similarity 95.1%; Pred. No. 8.8e-48;
 Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 ACTTCATTCTTCAAGTACAGACAGTGTGTCGTCCTGTTGCTAGCTACTTAAGAAG 60
 |||||
 DB 37 ACTTCATTCTTCAAGTACAGACAGTGTGTCGTCCTGTTGCTAGCTACTTAAGAAG 96

OY 61 CGGTGTTTTCAGAAATGATTAAGACCATGGCAAAATGTCAGCTTGGCTTGAAGA 120
 |||||
 DB 97 CGGTGTTTTCAGAAATGATTAAGACCATGGCAAAATGTCAGCTTGGCTTGAAGA 156

OY 121 ACCCTGTGTGCGGTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTAGTAGTAAGA 180
 |||||
 DB 157 ACCCTGTGTGCGGTTCTATTGCACAGAAATCAGAGCCCTCATTCCTTAGTAGTAAGA 216

```
RESULT 7
ID AAA94249 standard; DNA: 720 BP.
XX
AC AAA94249;
XX
DT 23-FEB-2001 (first entry)
XX
DE Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 5.
XX
KW Human; caspase activator; Smac; hybridising sequence; apoptosis; cancer;
KM autoimmune disease; neurodegenerative disease; mitochondria; ss.
XX
OS Homo sapiens.
XX
PN US6110691-A.
XX
PD 29-AUG-2000.
XX
PF 06-JAN-2000; 2000US-0479309.
XX
PR 06-JAN-2000; 2000US-0479309.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Wang X, Du C;
XX
DR WPI; 2000-586350/55.
XX
PT Novel caspase regulatory polypeptides useful for screening binding
PT agents specific for the polypeptides which are useful for diagnosis and
PT also for treating apoptosis associated diseases
XX
PS Disclosure; column 25-28; 16pp; English.
XX
CC The present sequence is a sequence which was shown to hybridise to the
CC coding sequence of the human Smac protein. The Smac coding sequence
CC was isolated by purifying the protein and searching a HeLa cell cDNA
CC library for sequences which bound to probes based upon it. Smac is a
CC mitochondrial protein which is released into the cytosol during
CC apoptosis, and acts as a caspase-3 activator. The coding sequence and
CC protein can be used to modulate the expression and function of caspases
CC and their activators, and also can be used as drug targets and regulators
CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
CC and neurodegenerative diseases.
XX
SQ Sequence 720 BP; 205 A; 151 C; 188 G; 176 T; 0 other;

Query Match 92.2%; Score 169.6; DB 21; Length 720;
Best Local Similarity 95.1%; Pred. No. 8.8e-48;
Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTGCTTCTGTTGGCTACTTTAAGAAG 60
DB 37 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTGCTTCTGTTGGCTACTTTAAGAAG 96
QY 61 CGGTGTTTCTCAGATATGATTAAGACCATGACACAAAACCTGAGAGATGGCTTTGAGTA 120
DB 97 CGGTGTTTCTCAGATATGATTAAGACCATGAGCTCAAACTGTGAGCATAGCCTTTGAGTA 156
QY 121 ACCCTGTGTCGGTTCCTATTGACACAGAAATCAGAGCCTCATTCCTTAGTAGAAGCA 180
DB 157 ACCCTGTGTCGGTTCCTATTGACAGTAATCAGAGCCTCATTCCTTAGTAGAAGCA 216
QY 181 TTGA 184
DB 217 TTGA 220

RESULT 8
ID AAA94250
```

```
ID AAA94250 standard; DNA: 720 BP.
XX
AC AAA94250;
XX
DT 23-FEB-2001 (first entry)
XX
DE Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 6.
XX
KW Human; caspase activator; Smac; hybridising sequence; apoptosis; cancer;
KM autoimmune disease; neurodegenerative disease; mitochondria; ss.
XX
OS Homo sapiens.
XX
PN US6110691-A.
XX
PD 29-AUG-2000.
XX
PF 06-JAN-2000; 2000US-0479309.
XX
PR 06-JAN-2000; 2000US-0479309.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Wang X, Du C;
XX
DR WPI; 2000-586350/55.
XX
PT Novel caspase regulatory polypeptides useful for screening binding
PT agents specific for the polypeptides which are useful for diagnosis and
PT also for treating apoptosis associated diseases
XX
PS Disclosure; column 27-28; 16pp; English.
XX
CC The present sequence is a sequence which was shown to hybridise to the
CC coding sequence of the human Smac protein. The Smac coding sequence
CC was isolated by purifying the protein and searching a HeLa cell cDNA
CC library for sequences which bound to probes based upon it. Smac is a
CC mitochondrial protein which is released into the cytosol during
CC apoptosis, and acts as a caspase-3 activator. The coding sequence and
CC protein can be used to modulate the expression and function of caspases
CC and their activators, and also can be used as drug targets and regulators
CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
CC and neurodegenerative diseases.
XX
SQ Sequence 720 BP; 208 A; 143 C; 193 G; 176 T; 0 other;

Query Match 84.3%; Score 155.2; DB 21; Length 720;
Best Local Similarity 90.2%; Pred. No. 7.2e-43;
Matches 166; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTGCTTCTGTTGGCTACTTTAAGAAG 60
DB 37 ACTTCATTTCTTCAGTACAGACAGTGTGTTGTTGCTTCTGTTGGCTACTTTAAGAAG 96
QY 61 CGGTGTTTCTCAGATATGATTAAGACCATGACACAAAACCTGAGAGATGGCTTTGAGTA 120
DB 97 CGGTGTTTCTCAGATATGATTAAGACCATGAGCTCAAACTGTGAGCATAGCCTTTGAGTA 156
QY 121 ACCCTGTGTCGGTTCCTATTGACACAGAAATCAGAGCCTCATTCCTTAGTAGAAGCA 180
DB 157 ACCCTGTGTCGGTTCCTATTGCTCTGAATACAGAGCCTCATTCCTTAGTAGAAGCA 216
QY 181 TTGA 184
DB 217 TTGA 220

RESULT 9
ID AAA95001 standard; DNA: 720 BP.
XX
AC AAA95001;
XX
```

DT 23-FEB-2001 (first entry)
XX
DE Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 7.
XX
KW Human; caspase activator; Smac; hybridising sequence; apoptosis; cancer;
XX autoimmune disease; neurodegenerative disease; mitochondria; ss.
OS Homo sapiens.
XX
PN US6110691-A.
XX
PD 29-AUG-2000.
XX
PE 06-JAN-2000; 2000US-0479309.
XX
PR 06-JAN-2000; 2000US-0479309.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Wang X, Du C;
XX
DR WPI; 2000-586350/55.
XX
PT Novel caspase regulatory polypeptides useful for screening binding
PT agents specific for the polypeptides which are useful for diagnosis and
PT also for treating apoptosis associated diseases -
XX
PS Disclosure; column 27-28; 16pp; English.
XX
CC The present sequence is a sequence which was shown to hybridise to the
CC coding sequence of the human Smac protein. The Smac coding sequence
CC was isolated by purifying the protein and searching a HeLa cell cDNA
CC library for sequences which bound to probes based upon it. Smac is a
CC mitochondrial protein which is released into the cytosol during
CC apoptosis, and acts as a caspase-3 activator. The coding sequence and
CC protein can be used to modulate the expression and function of caspases
CC and their activators, and also can be used as drug targets and regulators
CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
CC and neurodegenerative diseases.
XX
SQ Sequence 720 BP; 199 A; 162 C; 191 G; 168 T; 0 other;
Query Match 84.3%; Score 155.2; DB 21; Length 720;
Best Local Similarity 90.2%; Pred. No. 7.2e-43;
Matches 166; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTGTCCTCTGTGCTACTTAAAG 60
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
37 ACTTCATTTCTTCAGTACAGACAGTGTGTGTCCTCTGTGCTACTTAAAG 96
QY 61 CGGTGTTCTCAGAAATTGATAGACCATGCGACAAACTGTGAGATTGGAGTA 120
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
97 CGGTGTTCTCAGAAATTGATAGACCATGCGACAAACTGTGAGATTGGAGTA 156
QY 121 ACCCTGTGCGGTCCTATGTCAGCAAAATCAGACCTTCATTCCTAGTAGTGA 180
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
157 ACCCTGTGCGGTCCTATGTCAGCAAAATGAGGACCTCATTCCTAGTAGTGA 216
QY 181 TTGA 184
DB ||||
217 TTGA 220
RESULT 10
ID AAF66561
XX AAF66561 standard; cDNA; 376 BP.
XX
AC AAF66561;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 2317.
XX

KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PE 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
XX
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 883; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 376 BP; 105 A; 70 C; 92 G; 108 T; 1 other;
Query Match 83.7%; Score 154; DB 22; Length 376;
Best Local Similarity 99.4%; Pred. No. 1.5e-42;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 30 GTGTGTTCTGTTGTGTCGAATTAAGAGCGGTGTTCTCGAATTGTAAGAC 89
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
9 GTGTGTTCTGTTGTGTCGAATTAAGAGCGGTGTTCTCGAATTGTAAGAC 68
QY 90 GCACAAACTGTGAGCATTTGGCTTTGAGTAACCTGTGTGCGTCTTATTTGCACAGAA 149
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
69 GCACAAACTGTGAGCATTTGGCTTTGAGTAACCTGTGTGCGTCTTATTTGCACAGAA 128
QY 150 ATCAGAGCCTCAATTCCTTAGTAGTAAGACATTGA 184
DB ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
129 ATCAGAGCCTCAATTCCTTAGTAGTAAGACATTGA 163
RESULT 11
ID AAA95002
XX AAA95002 standard; DNA; 720 BP.
XX
AC AAA95002;
XX
DT 23-FEB-2001 (first entry)
XX

DE Human caspase activator Smac cDNA hybridising sequence SEQ ID NO: 8.
XX
XX Human: caspase activator: Smac; hybridising sequence; apoptosis; cancer;
KW autoimmune disease; neurodegenerative disease; mitochondria; ss.
XX
OS Homo sapiens.
XX
XX US6110691-A.
XX
XX 29-AUG-2000.
XX
XX 06-JAN-2000; 2000US-0479309.
XX
XX 06-JAN-2000; 2000US-0479309.
XX
XX (TEXA) UNIV TEXAS SYSTEM.
XX
XX Wang X, Du C;
XX
XX WPI: 2000-586350/55.
XX
XX
XX Novel caspase regulatory polypeptides useful for screening binding
PT agents specific for the polypeptides which are useful for diagnosis and
PT also for treating apoptosis associated diseases -
XX
XX
XX Disclosure; column 29-30; 16pp: English.
XX
XX The present sequence is a sequence which was shown to hybridise to the
CC coding sequence of the human Smac protein. The Smac coding sequence
CC was isolated by purifying the protein and searching a HeLa cell cDNA
CC library for sequences which bound to probes based upon it. Smac is a
CC mitochondrial protein which is released into the cytosol during
CC apoptosis, and acts as a caspase-3 activator. The coding sequence and
CC protein can be used to modulate the expression and function of caspases
CC and their activators, and also can be used as drug targets and regulators
CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
CC and neurodegenerative diseases.
XX
XX Sequence 720 BP; 218 A; 150 C; 184 G; 168 T; 0 other;
SQ
Query Match 83.3%; Score 153.2; DB 21; Length 720;
Best Local Similarity 90.1%; Pred. No. 3.5e-42;
Matches 164; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

XX
XX Homo sapiens.
OS
XX
XX WO200102568-A2.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18374.
XX
XX 02-JUL-1999; 99US-0142310.
XX
XX 02-JUL-1999; 99US-0142311.
XX
XX (CHIR) CHIRON CORP.
XX
XX (HYSE-) HYSEO INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamsom G, Dirmacac R;
PI Cikenjakov R, Dirmacac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
XX WPI: 2001-091805/10.
XX
XX
XX Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
XX
XX Claim 9; Page 654; 1046pp: English.
XX
XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as generic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
XX Sequence 432 BP; 118 A; 84 C; 104 G; 126 T; 0 other;
SQ
Query Match 81.7%; Score 150.4; DB 22; Length 432;
Best Local Similarity 88.6%; Pred. No. 2.6e-41;
Matches 163; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

DT 09-APR-2001 (first entry)
XX Novel human polynucleotide, SEQ ID NO: 2304.
XX
XX
XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
XX breast cancer; lung cancer; cancer detection; ss.
XX
XX Homo sapiens.
XX
XX WO200102568-A2.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18374.
XX
XX 02-JUL-1999; 99US-0142310.
XX 02-JUL-1999; 99US-0142311.
XX
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
XX WPI: 2001-091805/10.
XX
XX Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
XX Claim 9; Page 881-882; 1046pp; English.
XX
XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
XX Sequence 383 BP; 108 A; 72 C; 90 G; 113 T; 0 other;
XX
XX Query Match 80.8%; Score 148.6; DB 22; Length 383;
Best Local Similarity 97.4%; Pred. No. 1e-40;
Matches 151; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 30 GTGTTCTCTGTTGGCTTAAGAGGGGTTCTCAGAAATGATGAAGCCATG 89
DB 9 GTGTCTCTGTGTGGCTATTTTAAAGATGGGTCTTCAGATTGATGAAGCCATG 68
XX
XX 90 GCACAAACTGTGACGATTGGCTTGGAGTAACCTGTGTGGGTTCTTATTCACAGAA 149
DB 69 GCACAAACTGTGACGATTGGCTTGGAGTAACCTGTGTGGGTTCTTATTCACAGAA 128
XX
XX 150 ATCAGAGCCTCATTCCTTAACTAGTGAAGCATTTGA 184
DB 129 ATCAGAGCCTCATTCCTTAACTAGTGAAGCATTTGA 163
XX
XX RESULT 14
AAF64278
ID AAF64278 standard; cDNA: 227 BP.
XX

AC AAF64278;
XX
XX 09-APR-2001 (first entry)
XX
XX Novel human polynucleotide, SEQ ID NO: 34.
XX
XX
XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
XX breast cancer; lung cancer; cancer detection; ss.
XX
XX Homo sapiens.
XX
XX WO200102568-A2.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18374.
XX
XX 02-JUL-1999; 99US-0142310.
XX 02-JUL-1999; 99US-0142311.
XX
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
XX WPI: 2001-091805/10.
XX
XX Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
XX Claim 9; Page 549; 1046pp; English.
XX
XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
XX Sequence 227 BP; 42 A; 60 C; 68 G; 57 T; 0 other;
XX
XX Query Match 52.7%; Score 97; DB 22; Length 227;
Best Local Similarity 100.0%; Pred. No. 3.4e-23;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ACTTCATCTCTCAGGTACAGACAGTGTGTGTCTCTGTGTGGCTAACTTAAAG 60
DB 131 ACTTCATCTCTCAGGTACAGACAGTGTGTGTCTCTGTGTGGCTAACTTAAAG 190
XX
XX 61 CGGTGTTCTCAGAAATGATGAAGACATGACACAAA 97
DB 191 CGGTGTTCTCAGAAATGATGAAGACATGACACAAA 227
XX
XX RESULT 15
AAH14258
ID AAH14258 standard; cDNA: 2361 BP.
XX
XX AAH14258;
XX

```

DT 26-JUN-2001 (first entry)
XX
XX
DE Human cDNA sequence SEQ ID NO:11569.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX EPI074617-A2.
PN
XX
XX 07-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99UP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX
XX Claim 8; SEQ ID 11569; 2537bp + CD ROM; English.
XX
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX
SQ Sequence 2361 BP; 517 A; 617 C; 627 G; 600 T; 0 other;
XX
XX
XX Query Match 21.7%; Score 40; DB 22; Length 2361;
XX Best Local Similarity 100.0%; Pred. No. 0.0022;
XX Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 145 CAGAAATCAGAGCCTCATTCCTTAGTGAAGCATTTGA 184
DB 1208 CAGAAATCAGAGCCTCATTCCTTAGTGAAGCATTTGA 1247

```

"S PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 02:29:44 ; Search time 30244 Seconds

(without alignments)
177.057 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 184

Sequence: 1 attcattctctcagtgacag.....ccttagtagtgaagcattga 184

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_or:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgt_mus:*
34: em_hgt_pln:*
35: em_hgt_rtd:*
36: em_hgt_mam:*
37: em_hgt_vrt:*
38: em_sy:*
39: em_hgtgo_hum:*
40: em_hgtgo_mus:*
41: em_hgtgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------------|
| 1 | 184 | 100.0 | 720 | 6 | ARI07983 Sequence |
| 2 | 184 | 100.0 | 720 | 6 | AX191368 Sequence |
| 3 | 184 | 100.0 | 1358 | 6 | AX392556 Sequence |
| 4 | 184 | 100.0 | 1358 | 9 | AF262240 Homo sapi |
| 5 | 184 | 100.0 | 1371 | 9 | BC011909 Homo sapi |
| 6 | 182.4 | 99.1 | 1374 | 9 | AK024768 Homo sapi |
| 7 | 169.6 | 92.2 | 720 | 6 | ARI07984 Sequence |
| 8 | 169.6 | 92.2 | 720 | 6 | ARI07985 Sequence |
| 9 | 169.6 | 92.2 | 720 | 6 | ARI07986 Sequence |
| 10 | 169.6 | 92.2 | 720 | 6 | AX191370 Sequence |
| 11 | 169.6 | 92.2 | 720 | 6 | AX191371 Sequence |
| 12 | 169.6 | 92.2 | 720 | 6 | AX191372 Sequence |
| 13 | 155.2 | 84.3 | 720 | 6 | ARI07987 Sequence |
| 14 | 155.2 | 84.3 | 720 | 6 | ARI07988 Sequence |
| 15 | 155.2 | 84.3 | 720 | 6 | AX191373 Sequence |
| 16 | 155.2 | 84.3 | 720 | 6 | AX191374 Sequence |
| 17 | 154 | 83.7 | 376 | 6 | AX071845 Sequence |
| 18 | 153.2 | 83.3 | 720 | 6 | ARI07989 Sequence |
| 19 | 153.2 | 83.3 | 720 | 6 | AX191375 Sequence |
| 20 | 150.4 | 81.7 | 432 | 6 | AX070289 Sequence |
| 21 | 148.6 | 80.8 | 383 | 6 | AX071832 Sequence |
| 22 | 145.4 | 79.0 | 1327 | 9 | AK057778 Homo sapi |
| 23 | 139 | 75.5 | 189847 | 2 | AC048338 Homo sapi |
| 24 | 132 | 75.5 | 323263 | 2 | AC079406 Homo sapi |
| 25 | 122 | 66.3 | 1356 | 10 | AF203914 Mus muscu |
| 26 | 120.4 | 65.4 | 2585 | 10 | BC024780 Mus muscu |
| 27 | 97 | 52.7 | 227 | 6 | AX069562 Sequence |
| 28 | 94.6 | 51.4 | 224847 | 2 | AC129569 Mus muscu |
| 29 | 61 | 33.2 | 61784 | 2 | AC094200 Rattus no |
| 30 | 53 | 28.8 | 1320 | 9 | BC004417 Homo sapi |
| 31 | 40.8 | 22.2 | 189947 | 2 | AC048338 Homo sapi |
| 32 | 40 | 21.7 | 561 | 9 | AF298770 Homo sapi |
| 33 | 40 | 21.7 | 2361 | 9 | AK001399 Homo sapi |
| 34 | 39 | 21.2 | 2766 | 9 | AL833244 Homo sapi |
| 35 | 35 | 19.0 | 4009 | 1 | NOSORF M81381 Nostoc sp. |
| 36 | 34.2 | 18.6 | 129639 | 2 | AC116809 Mus muscu |
| 37 | 33.6 | 18.3 | 208160 | 2 | AC124254 Homo sapi |
| 38 | 33.4 | 18.2 | 153900 | 2 | AC116408 Mus muscu |
| 39 | 33.4 | 18.2 | 155925 | 9 | AC096536 Homo sapi |
| 40 | 33.4 | 18.2 | 169072 | 9 | AC009946 Homo sapi |
| 41 | 33.4 | 18.2 | 177089 | 9 | AC008277 Homo sapi |
| 42 | 33.4 | 18.2 | 182563 | 2 | AC129033 Rattus no |
| 43 | 33.2 | 18.0 | 143848 | 2 | AC126584 Rattus no |
| 44 | 33.2 | 18.0 | 183522 | 2 | AC107568 Rattus no |
| 45 | 33 | 17.9 | 56128 | 9 | AC005349 Homo sapi |

ALIGNMENTS

RESULT 1
LOCUS ARI07983 720 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6110691.
ACCESSION ARI07983
VERSION ARI07983.1 GI:12823470
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang, X. and Du, C.
TITLE Activators of caspases
JOURNAL Patent: US 6110691-A 1 29-AUG-2000;
FEATURES Location/Qualifiers

| | | |
|------------|---------------------|-------|
| source | 1. .720 | |
| | /organism="unknown" | |
| BASE COUNT | 206 a 147 c 195 g | 172 t |
| ORIGIN | | |

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 100.0% | Score 184; | DB 6; | Length 720; |
| Best Local Similarity | 100.0% | Pred. No. 2.1e-49; | | |
| Matches 184; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

[illegible]

| | | | | |
|------------|-----------------------------------|-------------|-----|--------|
| RESULT 2 | | | | |
| LOCUS | AX191368 | | | |
| DEFINITION | AX191368 | 720 bp | DNA | |
| ACCESSION | Sequence 1 from Patent WO0149719. | | | linear |
| VERSION | AX191368 | | | |
| KEYWORDS | AX191368.1 | GI:15209588 | | |
| SOURCE | human. | | | |

REFERENCE Eumariota, Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
 1 (bases 1 to 720)
AUTHORS Wang, X.U. and Du, C. U.
TITLE Activators of caspases
JOURNAL Patent: WO 0149719-A 1 12-JUL-2001;
FEATURES BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM (US)
 Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
l. .720
CDS

```

```

/note="unnamed protein product"
/codon_start=1
/protein_id="CAC5136.1"
/db_xref="GI:15209589"
/translation="MAALSKSLRSVTSFPRROCLCPVYANFKKRCFSLIRPMK
TWTGPGVTLCAVPIAKSPKSHSLKIMRAVSLVTDSTFSLQTYALIAIPI
YTRKYVTLTSLYROYTSLGLCKMNSDEEDYQVILGARAEKNTSHQELKLETTMNR
KVGSEEMAEAYVQTGAODASTTARNHHQLQVLAQVEEVHQLSRKAETKLAEQIIELR
KTEQGEGRARSEDEAVLRD"

```

| | | | | |
|---------------------------|--------|--------------------|-----------|-------------|
| Query Match | 100.0% | Score 184; | DB 6; | Length 720; |
| Best Local Similarity | 100.0% | Pred. No. 2.1e-49; | | |
| Matches 184; Conservative | 0; | Mismatches 0; | Indels 0; | Gaps 0; |

Oy 1 ACTTCAATCTTCAGGTACACAGACAGTGTGTGTGTCTCTGTTGGCTAACCTTTAAGAAG 60
 Db 37 ACTTCAATCTTCAGGTACAGACAGTGTGTGTGTCTCTGTTGGCTAACCTTTAAGAAG 96
 Oy 61 CGGTGTTCTCAGAAATGATATAAGACCATTGGCCAAACACTGATGACATTGGCTTGGAGTA 120
 Db 97 CGGTGTTCTCAGAAATGATATAAGACCATTGGCCAAACACTGATGACATTGGCTTGGAGTA 156
 Oy 121 ACCCTGTGTGGGTTCCATTTCACAGAAATGACAGCCTCATTTCCCTTAGTAGTAAGA 180

| | | | |
|----|-----|--|-----|
| Db | 157 | ACCTGTGTGGCGTTCCTATTGCACGAATCAAGGCTCATTTCCCTAGTAGTGAAGCA | 216 |
| Qy | 181 | TTGA | 184 |
| | | | |
| Db | 217 | TTGA | 220 |

| RESULT 3 | | | |
|------------|-----------------------------------|-------------|-----------------|
| AX392556 | | | |
| LOCUS | AX392556 | 1358 bp | DNA |
| DEFINITION | Sequence 1 from Patent WO0216418. | | Linear |
| ACCESSION | AX392556 | | PAT 23-MAR-2002 |
| VERSION | AX392556.1 | GI:19700793 | |

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

1
Alnemri, E. S.
An iap binding peptide or polypeptide and methods of using the same
Patent: WO 0216418-A 1 28-FEB-2002;
THOMAS JEFFERSON UNIVERSITY (US)
location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
20..739
CDS

```

/translation-"MAAIKMSLMSRSYSPFRRCGLCPVPAVNAIKKCESELLRPMK
 TYVTGQVATLCAPVIAKQSEBHSISALMARAVSLTDDTSFPLSOTYALLAEATE
 YTKAVTTLISLIRQYISLIGLPMNSDEEWMQYIGRAEETSCHQYLWLETTWTA
 VGSEMAEAAYQFGADQASITANRHQLVLQVLEEVOHLSRAETKLAIAQGLELKO
 KTOEGEGERASEDEAVLRD"
 341 a 314 c 358 g 345 t

| | | | | |
|---------------------------|--------|--------------------|-------|-------------------|
| Query Match | 100.0% | Score 184; | DB 6; | Length 1358; |
| Best Local Similarity | 100.0% | Pred. No. 2.1e-49; | | |
| Matches 184; Conservative | 0; | Mismatches | 0; | Indels 0; Gaps 0; |

| | | | |
|----|-----|--|-----|
| QY | 1 | ACTTATCTTCAGGTTACAGACAGCTTGTGTGTCTGTGTGGCTAACTTTAAGAG | 60 |
| Db | 56 | ACTTATCTTCAGGTTACAGACAGCTTGTGTGTCTGTGTGGCTAACTTTAAGAG | 115 |
| QY | 61 | CGGTGTTCTCAGATTGATTAAGACCATGGACAACAACGTGACAGATTGGCTTGAAGTA | 120 |
| Db | 116 | CGGTGTTCTCAGATTGATTAAGACCATGGACAACAACGTGACAGATTGGCTTGAAGTA | 175 |
| QY | 121 | ACCTGTGTGGGGTTCCTATTGACACAGAAATCAGAGCCTCATTCCTTTAGTAGTGAACGA | 180 |
| Db | 176 | ACCTGTGTGGGGTTCCTATTGACACAGAAATCAGAGCCTCATTCCTTTAGTAGTGAACGA | 235 |
| QY | 181 | TTTGA 184 | |
| Db | 236 | TTTGA 239 | |

| | | | | | |
|------------|---|------------|------|--------|-----------------|
| LOCUS | AF262240 | 1358 bp | mRNA | linear | PRI 26-JUL-2000 |
| DEFINITION | Homo sapiens Smac mRNA, complete cds; nuclear gene for mitochondrial product. | | | | |
| ACCESSION | AF262240 | | | | |
| VERSION | AF262240.1 | GI:9454218 | | | |
| KEYWORDS | | | | | |
| SOURCE | Homo sapiens. | | | | |
| ORGANISM | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |

REFERENCE 1 (bases 1 to 1358)
 Du.C., Fang, M., Li, Y., Li, L. and Wang, X.
 TITLE Smac, a mitochondrial protein that promotes cytochrome c-dependent
 caspase activation by eliminating IAP inhibition
 JOURNAL Cell 102 (1), 33-42 (2000)
 MEDLINE 20383536
 PUBMED 10929711
 REFERENCE 2 (bases 1 to 1358)
 Du.C., Fang, M., Li, Y. and Wang, X.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAY-2000) Howard Hughes Medical Institute and
 Department of Biochemistry, University of Texas Southwestern Medical
 Center, 5323 Harry Hines Blvd., Dallas, TX 75235, USA
 FEATURES
 source
 1..1358
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 20..739
 /function="binds IAPs and neutralizes their inhibition on
 caspase activation and activity"
 /note="antagonist of IAPs (inhibitors of apoptosis)"
 /codon_start=1
 /product="Smac"
 /protein_id="AA087716.1"
 /db_xref="GI:9454219"
 /translation="MAALKSWLSRSVTSFFRYROCLCVPVVAFKRCFSELIRPMHK
 TVTIGFVTLCAVPFAOKSEPHSLSEALMRVAVLVDSTSTFISOTTYALIEATE
 YTKAVYTLISLYROYTSLGKKNSEEDBNVQVITGARAEMTSKHQETLKLETTWMTA
 VGLSMAAEAAVOTGADQASITARNHIQVLKQVEVHOLSKKATKLAEOIEELRO
 KTOEGEERAESEDEAVYLRD"
 translt_peptide 20..184
 /note="mitochondrial targeting sequence"
 BASE COUNT 341 a 314 c 358 g 345 t
 ORIGIN
 Query Match 100.0%; Score 184; DB 9; Length 1358;
 Best Local Similarity 100.0%; Pred. No. 2,1e-49;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACTTCATTTCTCAGTACAGACAGTGTGTGTCCTGTTGCGCTAACTTTAAGAAG 60
 |||||||
 DB 56 ACTTCATTTCTCAGTACAGACAGTGTGTGTCCTGTTGCGCTAACTTTAAGAAG 115
 |||||||
 QY 61 CGGTGTTTCTCAGAAATTGATAGACCATGACAAACATGTCAGATTGGCTTTGAGTA 120
 |||||||
 DB 116 CGGTGTTTCTCAGAAATTGATAGACCATGACAAACATGTCAGATTGGCTTTGAGTA 175
 |||||||
 QY 121 ACCCTGTGTGCGGTTCTATTGACAGAAATCAGAGCTTCATTCCTTAGTAGTAAGA 180
 |||||||
 DB 176 ACCCTGTGTGCGGTTCTATTGACAGAAATCAGAGCTTCATTCCTTAGTAGTAAGA 235
 |||||||
 QY 181 TTGA 184
 |||||
 DB 236 TTGA 239
 |||||
 RESULT 5
 BC011909 1371 bp mRNA linear PRI 02-AUG-2001
 LOCUS Homo sapiens, Similar to second mitochondria-derived activator of
 DEFINITION caspase, clone MGC:19863 IMAGE:4137792, mRNA, complete cds.
 ACCESSION BC011909
 VERSION BC011909.1 GI:15080296
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1371)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Shenchenko, Y., Metherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, Q.L., Mastello, C., Mastriani, S.D., McCloskey, D.C.,
 McDowell, J., Pearson, R., Snyder, B., Stantirlop, S., Thomas, P.J.,
 Tlionson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 28 Row: k Column: 12
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10437143.
 FEATURES
 source
 1..1371
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:19863 IMAGE:4137792"
 /tissue_type="Muscle, rhabdomyosarcoma"
 /clone_lib="NHG-MGC_17"
 /lab_host="DH10B-R"
 /note="vector: pOTB7"
 33..752
 /codon_start=1
 /product="Similar to second mitochondria-derived activator
 of caspase"
 /protein_id="AAH11909.1"
 /db_xref="GI:15080297"
 /translation="MAALKSWLSRSVTSFFRYROCLCVPVVAFKRCFSELIRPMHK
 TVTIGFVTLCAVPFAOKSEPHSLSEALMRVAVLVDSTSTFISOTTYALIEATE
 YTKAVYTLISLYROYTSLGKKNSEEDBNVQVITGARAEMTSKHQETLKLETTWMTA
 VGLSMAAEAAVOTGADQASITARNHIQVLKQVEVHOLSKKATKLAEOIEELRO
 KTOEGEERAESEDEAVYLRD"
 BASE COUNT 349 a 314 c 363 g 345 t
 ORIGIN
 Query Match 100.0%; Score 184; DB 9; Length 1371;
 Best Local Similarity 100.0%; Pred. No. 2,1e-49;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACTTCATTTCTCAGTACAGACAGTGTGTGTCCTGTTGCGCTAACTTTAAGAAG 60
 |||||||
 DB 69 ACTTCATTTCTCAGTACAGACAGTGTGTGTCCTGTTGCGCTAACTTTAAGAAG 128
 |||||||
 QY 61 CGGTGTTTCTCAGAAATTGATAGACCATGACAAACATGTCAGATTGGCTTTGAGTA 120
 |||||||
 DB 129 CGGTGTTTCTCAGAAATTGATAGACCATGACAAACATGTCAGATTGGCTTTGAGTA 188
 |||||||
 QY 121 ACCCTGTGTGCGGTTCTATTGACAGAAATCAGAGCTTCATTCCTTAGTAGTAAGA 180
 |||||||
 DB 189 ACCCTGTGTGCGGTTCTATTGACAGAAATCAGAGCTTCATTCCTTAGTAGTAAGA 248
 |||||||
 QY 181 TTGA 184
 |||||
 DB 249 TTGA 252
 |||||
 RESULT 6
 AK024768 1374 bp mRNA linear PRI 29-SEP-2000
 LOCUS AK024768

```

DEFINITION Homo sapiens cDNA: FLJ21115 fis, clone CAS05491.
ACCESSION AK024768
VERSION AK024768.1 GI:10437143
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens primary smooth muscle cells of human coronary artery
          cDNA to mRNA, clone_11D:CAS clone:CAS05491.
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Watanabe,K., Kunagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
          Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
          Nakamura,Y., Isogai,T. and Sugano,S.
          NEDO human cDNA sequencing project
          Unpublished
          2 (bases 1 to 1374)
          Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
          Shibahara,T., Tanaka,T. and Nakamura,Y.
          Direct Submission
          Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
          University of Tokyo, Laboratory of Genome Structure Analysis, Human
          Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
          Japan (E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
          Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
          International Trade and Industry of Japan; cDNA full insert
          sequencing: Research Association for Biotechnology; cDNA library
          construction, 5'- 6' 3'-end one pass sequencing: Department of
          Virology and Human Genome Center, Institute of Medical Science,
          University of Tokyo (partly supported by Science and Technology
          Agency).
FEATURES
    source
        1..1374
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="CAS05491"
        /cell_type="primary smooth muscle cells of human coronary
        artery"
        /clone_11b="CAS"
        /note="cloning vector pME18SFL3"
        28..747
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="BAB14994.1"
        /db_xref="GI:10437144"
        /translation="MAALKSWLSRSVMSPEFYRCLOVPYVANKKRCSEFLIPMR
        YTKAVVLTSLYRQYISLGLKMSNEEEDVWQVLIIGAAEWTSKHQELKLETTWMTA
        VGLSEMAAEAAVYTGADQASITARNHLDLVKLOVEVHQLSRKAETKLAIEQIELRQ
        KTOEEGERASEBOEAYLRD"
    CDS
        351 a 314 c 361 g 348 t
        BASE COUNT
        ORIGIN
Query Match 99.1%; Score 182.4; DB 9; Length 1374;
Best Local Similarity 99.5%; Pred. No. 6.9e-49;
Matches 183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 ACTTATTCTTCAGGTACAGACAGTGTGTCCTGTCGTGGCTAACTTAAGAAG 60
    |||||||
DB 64 ACTTATTCTTCAGGTACAGACAGTGTGTCCTGTCGTGGCTAACTTAAGAAG 123
OY 61 CGGTGTTCTCAGAAATGATAGACATGGCAAAACTGTGACATGGCTTTGGAGTA 120
    |||||||
DB 124 CGGTGTTCTCAGAAATGATAGACATGGCAAAACTGTGACATGGCTTTGGAGTA 183
OY 121 ACCCTGTGCGGTCTCAATTCACAGAAATCAGAGCTCATTCCTTAGTAGTAAGCA 180
    |||||||
DB 184 ACCCTGTGCGGTCTCAATTCACAGAAATCAGAGCTCATTCCTTAGTAGTAAGCA 243
OY 181 TTGA 184
    |||||
DB 244 TTGA 247

```

```

RESULT 7
LOCUS AR107984
DEFINITION Sequence 3 from patent US 6110691.
ACCESSION AR107984
VERSION AR107984.1 GI:12823471
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang,X. and Du,C.
TITLE Activators of caspases
JOURNAL Patent: US 6110691-A 3 29-AUG-2000;
FEATURES
    source
        1..720
        /organism="unknown"
    BASE COUNT
    ORIGIN
Query Match 92.2%; Score 169.6; DB 6; Length 720;
Best Local Similarity 95.1%; Pred. No. 1.1e-44;
Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
OY 1 ACTTCATCTTCAGGTACAGACAGTGTGTCCTGTCGTGGCTAACTTAAGAAG 60
    |||||||
DB 37 ACTTCATCTTCAGGTACAGACAGTGTGTCCTGTCGTGGCTAACTTAAGAAG 96
OY 61 CGGTGTTCTCAGAAATGATAGACATGGCAAAACTGTGACATGGCTTTGGAGTA 120
    |||||||
DB 97 CGGTGTTCTCAGAAATGATAGACATGGCAAAACTGTGACATGGCTTTGGAGTA 156
OY 121 ACCCTGTGCGGTCTCAATTCACAGAAATCAGAGCTCATTCCTTAGTAGTAAGCA 180
    |||||||
DB 157 ACCCTGTGCGGTCTCAATTCACAGAAATCAGAGCTCATTCCTTAGTAGTAAGCA 216
OY 181 TTGA 184
    |||||
DB 217 TTGA 220
RESULT 8
LOCUS AR107985
DEFINITION Sequence 4 from patent US 6110691.
ACCESSION AR107985
VERSION AR107985.1 GI:12823472
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang,X. and Du,C.
TITLE Activators of caspases
JOURNAL Patent: US 6110691-A 4 29-AUG-2000;
FEATURES
    source
        1..720
        /organism="unknown"
    BASE COUNT
    ORIGIN
Query Match 92.2%; Score 169.6; DB 6; Length 720;
Best Local Similarity 95.1%; Pred. No. 1.1e-44;
Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
OY 1 ACTTCATCTTCAGGTACAGACAGTGTGTCCTGTCGTGGCTAACTTAAGAAG 60
    |||||||
DB 37 ACTTCATCTTCAGGTACAGACAGTGTGTCCTGTCGTGGCTAACTTAAGAAG 96
OY 61 CGGTGTTCTCAGAAATGATAGACATGGCAAAACTGTGACATGGCTTTGGAGTA 120
    |||||||
DB 97 CGGTGTTCTCAGAAATGATAGACATGGCAAAACTGTGACATGGCTTTGGAGTA 156

```

QY 121 ACCCTGTGCGGTTCTCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGGAACA 180
|||||
Db 157 ACCCTGTGCGGTTCTCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGGAACA 216
|||||

QY 181 TTGA 184
|||||
Db 217 TTGA 220

RESULT 9
ARI07986 720 bp DNA linear PAT 14-FEB-2001
LOCUS
DEFINITION Sequence 5 from patent US 6110691.
ACCESSION ARI07986
VERSION ARI07986.1 GI:12823473
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang, X. and Du, C.
TITLE Activators of caspases
JOURNAL Patent: US 6110691-A 5 29-AUG-2000;
FEATURES
source 1..720
location/Qualifiers
BASE COUNT 205 a 151 c 188 g 176 t
ORIGIN

Query Match 92.2%; Score 169.6; DB 6; Length 720;
Best Local Similarity 95.1%; Pred. No. 1.1e-44;
Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTGTTCTCTGTTGGCTTAAGTAAAG 60
|||||
Db 37 ACTTCATTTCTTCAGTACAGACAGTGTGTGTTCTCTGTTGGCTTAAGTAAAG 96
|||||

QY 61 CGGTTTCTCAGAAATGATAGAACCACTGACAAACTGACAGATTGGCTTTGGAGTA 120
|||||
Db 97 CGGTTTCTCAGAAATGATAGAACCACTGACAAACTGACAGATTGGCTTTGGAGTA 156
|||||

QY 121 ACCCTGTGCGGTTCTCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGGAACA 180
|||||
Db 157 ACCCTGTGCGGTTCTCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGGAACA 216
|||||

QY 181 TTGA 184
|||||
Db 217 TTGA 220

RESULT 10
AX191370 720 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 3 from Patent WO0149719.
ACCESSION AX191370
VERSION AX191370.1 GI:15209590
KEYWORDS
ORGANISM synthetic construct.
SOURCE synthetic construct.
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang, X.U. and Du, C.U.
TITLE Activators of caspases
JOURNAL Patent: WO 0149719-A 3 12-JUL-2001;
FEATURES
location/Qualifiers
source 1..720
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Sequence"
BASE COUNT 207 a 145 c 190 g 178 t
ORIGIN

Query Match 92.2%; Score 169.6; DB 6; Length 720;
Best Local Similarity 95.1%; Pred. No. 1.1e-44;
Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTGTTCTCTGTTGGCTTAAGTAAAG 60
|||||
Db 37 ACTTCATTTCTTCAGTACAGACAGTGTGTGTTCTCTGTTGGCTTAAGTAAAG 96
|||||

QY 61 CGGTTTCTCAGAAATGATAGAACCACTGACAAACTGACAGATTGGCTTTGGAGTA 120
|||||
Db 97 CGGTTTCTCAGAAATGATAGAACCACTGACAAACTGACAGATTGGCTTTGGAGTA 156
|||||

QY 121 ACCCTGTGCGGTTCTCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGGAACA 180
|||||
Db 157 ACCCTGTGCGGTTCTCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGGAACA 216
|||||

QY 181 TTGA 184
|||||
Db 217 TTGA 220

RESULT 11
AX191371 720 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 4 from Patent WO0149719.
ACCESSION AX191371
VERSION AX191371.1 GI:15209591
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 720)
AUTHORS Wang, X.U. and Du, C.U.
TITLE Activators of caspases
JOURNAL Patent: WO 0149719-A 4 12-JUL-2001;
FEATURES
location/Qualifiers
source 1..720
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic Sequence"
BASE COUNT 202 a 154 c 192 g 172 t
ORIGIN

Query Match 92.2%; Score 169.6; DB 6; Length 720;
Best Local Similarity 95.1%; Pred. No. 1.1e-44;
Matches 175; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTGTTCTCTGTTGGCTTAAGTAAAG 60
|||||
Db 37 ACTTCATTTCTTCAGTACAGACAGTGTGTGTTCTCTGTTGGCTTAAGTAAAG 96
|||||

QY 61 CGGTTTCTCAGAAATGATAGAACCACTGACAAACTGACAGATTGGCTTTGGAGTA 120
|||||
Db 97 CGGTTTCTCAGAAATGATAGAACCACTGACAAACTGACAGATTGGCTTTGGAGTA 156
|||||

QY 121 ACCCTGTGCGGTTCTCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGGAACA 180
|||||
Db 157 ACCCTGTGCGGTTCTCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGGAACA 216
|||||

QY 181 TTGA 184
|||||
Db 217 TTGA 220

RESULT 12
AX191372 720 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0149719.
ACCESSION AX191372
VERSION AX191372.1 GI:15209592
KEYWORDS


```
QY 1 ACTTCATTCTTCAGGTACAGACAGTGTGTTGCTCCTGTTGGCTAACTTAAGAG 60
    ||||||| ||||| ||| ||||||| ||||| ||| |||
Db 37 ACTTCATTCTTTAGTACTGACAGAGTTAGTGTCTCTGTTGGCTATCTTTAGAG 96
    ||||||| ||| ||||||| ||||||| ||||| |||
QY 61 CGGTGTTTCTCAGATTGATTAGACCAATGSCACAAAACGTGACGATTGGCTTTGAGTA 120
    ||||||| || ||||||| ||||||| ||||| |||
Db 97 CGGTGTTTCTCAAAATGATTAAGACCAAGCCACAAAACAGTGACGTTGGCTTTGAGTT 156
    ||||||| || ||||||| ||||||| ||||| |||
QY 121 ACCCTGTGTGGGTTCCCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGTGAAGCA 180
    ||||||| ||||| ||||||| ||||||| ||||| |||
Db 157 ACCCTGAGTGGCGATCCTATTGCTGTGAATCAGAGCCTCATTCCTAAGTAGAAGCA 216
    ||||||| ||||| ||||||| ||||||| ||||| |||
QY 181 TTGA 184
    |||
Db 217 TTGA 220
```

Search completed: February 16, 2003, 00:27:30
Job time : 30247 secs

THIS PAGE BLANK (USPTO)

QY 121 ACCCTGTGCGGTTCTATTGACAGAAATCAGAGCCCTCATTCCTTAGTAGTGAAGCA 180
|||||
Db 53 ThrleuGysAlaValProIleAlaGlnLysSerGluProHisSerLeuSerSerGluAla 72
QY 181 TTG 183
|||
Db 73 Leu 73
RESULT 2
US-09-134-001C-3138
; Sequence 3138, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3138
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3138
Alignment Scores:
Pred. No.: 1.04 Length: 190
Score: 62.00 Matches: 12
Percent Similarity: 72.00% Conservative: 6
Best Local Similarity: 48.00% Mismatches: 7
Query Match: 18.96% Indels: 0
Gaps: 0
Db: 4
US-09-939-293-1_COPY_56_239 (1-184) x US-09-134-001C-3138 (1-190)
QY 10 TTCAGGTACAGACAGCTTTGTGTGCTTCTGTGCGCTAACTTTAAGAGCGGTGTTTC 69
|||||
Db 2 TyrgInTyAspLeuAsnLeuLeuProLeuValLysIleLeuLysAsrGysPhe 21
QY 70 TCAGATTGATAGA 84
|||||
Db 22 SerLysMetIleArg 26
RESULT 3
US-09-134-001C-5353
; Sequence 5353, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5353
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5353
Alignment Scores:
Pred. No.: 8.11 Length: 121

Score: 55.50 Matches: 15
Percent Similarity: 51.11% Conservative: 8
Best Local Similarity: 33.33% Mismatches: 7
Query Match: 17.29% Indels: 15
Gaps: 3
US-09-939-293-1_COPY_56_239 (1-184) x US-09-134-001C-5353 (1-121)
QY 134 ACCGACACAGGAGGTACTTCACAAAGCCATCGTACAGATTTC-----TGCCATGCTCT 81
||| |||||
Db 82 ThrLeuHisLysValSerGlnLysLeuIleLysAsnIleIleLysTyThrTyrlGlu 101
QY 80 ---ATCAATTCGAGAAACACCGCTTCTTAAGTTAGCCACAAAGAAACACAAACAC 24
|||||
Db 102 GluIleAsnLysAspLys-----LysHis 109
QY 23 TGTCTGTACTTGAAG 9
|||||
Db 110 CysMetTyrlIleLys 114
RESULT 4
5378464-3
; Patent No. 5378464
; APPLICANT: MCEVER, RODGER P.
; TITLE OF INVENTION: MODULATION OF INFLAMMATORY RESPONSES
; BY ADMINISTRATION OF GMP-140 OR ANTIBODY TO GMP-140
; NUMBER OF SEQUENCES: 32
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/320,408
; FILING DATE: 08-MAR-1989
; SEQ ID NO: 3
; LENGTH: 574
5378464-3
Alignment Scores:
Pred. No.: 11.5 Length: 574
Score: 55.50 Matches: 11
Percent Similarity: 54.84% Conservative: 6
Best Local Similarity: 35.48% Mismatches: 9
Query Match: 16.97% Indels: 5
Gaps: 2
US-09-939-293-1_COPY_56_239 (1-184) x 5378464-3 (1-574)
QY 69 CTCAGAAATGATTAAGACCATGGCACAACCTGTGACGATTGGCTTGAGTAAC----- 122
||| |||||
Db 509 LeuHisValAsnLysProIleAlaMetAsnLysSerAsn---LeuTrpIleYasnPheSer 527
QY 123 -----CCGTGTGCGGTTCTATTGACAGAA 149
|||||
Db 528 TyrgLysSerIleCysSerPheHisCysLeuGlu 538
RESULT 5
US-09-134-001C-3363
; Sequence 3363, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3363
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3363


```

FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 163..174
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 168..183
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 185..194
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 200..244
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 213..226
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 230..257
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 262..306
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 275..288
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 292..319
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 324..368
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 337..350
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 354..381
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 386..430
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 399..412
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 416..443
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 448..492
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 461..474
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 478..505
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 510..554
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 523..536
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 540..567
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 572..616
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 585..616
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 602..629
FEATURE:

NAME/KEY: Disulfide-bond
LOCATION: 642..686
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 655..668
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 672..699
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 704..748
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 717..730
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: 734..761
FEATURE:
NAME/KEY: Binding-site
LOCATION: 54
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 98
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 180
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 212
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 219
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 411
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 460
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 518
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 665
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 716
OTHER INFORMATION: /note="Potential asparagine-linked
OTHER INFORMATION: glycosylation site"
FEATURE:

Alignment Scores:
Pred. No.: 12.5 Length: 830
Score: 55.50 Matches: 11
Percent Similarity: 54.84% Conservative: 6

```

```

1 NUMBER OF SEQUENCES: 112
2
3 CORRESPONDENCE ADDRESS:
4
5 ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
6
7 STREET: Four Embarcadero Center, Suite 3400
8
9 CITY: San Francisco
10
11 STATE: California
12
13 COUNTRY: US
14
15 ZIP: 94111
16
17
18 COMPUTER READABLE FORM:
19
20 MEDIUM TYPE: Floppy disk
21
22 COMPUTER: IBM PC compatible
23
24 OPERATING SYSTEM: PC-DOS/MS-DOS
25
26 SOFTWARE: Patentln Release #1.0, Version #1.30
27
28 CURRENT APPLICATION DATA:
29
30 APPLICATION NUMBER: US/08/482,085B
31
32 FILING DATE: 07-JUN-1995
33
34 CLASSIFICATION: 435
35
36 PRIOR APPLICATION DATA:
37
38 APPLICATION NUMBER: US 06/927,258
39
40 FILING DATE: 04-NOV-1986
41
42 PRIOR APPLICATION DATA:
43
44 APPLICATION NUMBER: US 07/114,618
45
46 FILING DATE: 29-OCT-1987
47
48 PRIOR APPLICATION DATA:
49
50 APPLICATION NUMBER: US 08/053,049
51
52 FILING DATE: 22-APR-1993
53
54 PRIOR APPLICATION DATA:
55
56 APPLICATION NUMBER: US 08/175,155
57
58 FILING DATE: 29-DEC-1993
59
60 ATTORNEY/AGENT INFORMATION:
61
62 NAME: Treccartin, Richard F.
63
64 REGISTRATION NUMBER: 31,801
65
66 REFERENCE/DOCKET NUMBER: A-55186-6/RFT/MTK
67
68 TELECOMMUNICATION INFORMATION:
69
70 TELEPHONE: 415-781-1989
71
72 TELEFAX: 415-398-3249
73
74 INFORMATION FOR SEQ ID NO: 20:
75
76 SEQUENCE CHARACTERISTICS:
77
78 LENGTH: 285 amino acids
79
80 TYPE: amino acid
81
82 STRANDEDNESS: single
83
84 TOPOLOGY: linear
85
86 MOLECULE TYPE: peptide
87
88 US-08-482-085B-20
89
90
91 Alignment Scores:
92
93 Pred. No.: 16.2 Length: 285
94
95 Score: 54.00 Matches: 13
96
97 Percent Similarity: 41.18% Conservative: 8
98
99 Best local Similarity: 25.49% Mismatches: 19
100
101 Query Match: 16.51% Indels: 11
102
103 Gaps: 3 1
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
```


Db 223 ProTyrHisAsnAlaValHisAlaAlaAspValThrGlnAlaMetHisCys---TyrIleu 241
QY 11 AAG 9
Db 242 Lys 242

RESULT 13
US-08-948-564-18
; Sequence 18, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminszky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: NO. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-948-564-18

Alignment Scores:
Pred. No.: 21.8 Length: 509
Score: 53.50 Matches: 18
Percent Similarity: 47.27% Conservative: 8
Best Local Similarity: 32.73% Mismatches: 18
Query Match: 16.36% Indels: 11
DB: 3 Gaps: 3

US-09-939-293-1_COPY_56_239 (1-184) x US-08-948-564-18 (1-509)

QY 37 CCGTGTGGTGAAGTCTT-----AAGAACGGGTGTTTCAGAAATG 78
Db 36 ProValValGlyAsnLeuTyrAspIleLysProValAlaArgPheArgCysPheAlaGluTrrp 55
QY 79 ATAAGACCATGGCACAACAACTGTGAGATGGCTTTGGAGTAACCCCTGTGCGGTTCCT 138
Db 56 AlaGlnSerTyrGlyProIleIleSerValAlaTrpPheGlySerThrLeuAsnVal----- 73

QY 139 ATGACAGAAATCAGAGCCTATTCCTTACTAGTGAAGCATTTG 183
Db 74 IleValSerAsnSerGlu-----LeuAlaLysGluValIleu 85

RESULT 14
US-08-356-786-10

; Sequence 10, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-786-10

Alignment Scores:
Pred. No.: 22 Length: 534
Score: 53.50 Matches: 17
Percent Similarity: 41.07% Conservative: 6
Best Local Similarity: 30.36% Mismatches: 22
Query Match: 16.36% Indels: 11
DB: 2 Gaps: 2

US-09-939-293-1_COPY_56_239 (1-184) x US-08-356-786-10 (1-534)

QY 1 ACTCATCTTCAGTACAGACAGAGTGTGTGTCGTGGCTAGCACTTAAGAAG 60
Db 241 SerLysPheSerValIleAspValSerIleIleuIleProIleIleAlaLeuMetValIle 260

QY 61 CGGTGT-----TTCTCAGAAATTGATTAAGACCATGGCAAAACT 99
Db 261 ArgCysAlaProProProSerSerGlnPheSerIleuIleAlaArgPro----- 276

QY 100 GTGACGATTTGCTTTGGAGTAACCTGTGTGCGGTTCTATTGCACAG 147
Db 277 ValValProAsnPheAsnAlaAspValCysMetAspProGluIleGln 292

RESULT 15
US-08-089-458B-6
; Sequence 6, Application US/08089458B
; Patent No. 5359039
; GENERAL INFORMATION:
; APPLICANT: Smith, Craig
; APPLICANT: Goodwin, Raymond

TITLE OF INVENTION: Isolated Poxvirus A53R-Equivalent Tumor
TITLE OF INVENTION: Necrosis Factor Antagonists
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patricia Anne Perkins, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System 7.1
SOFTWARE: Microsoft Word, Version #5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/089,458B
FILING DATE: 07/09/93
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 387-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 186 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-089-458B-6

Alignment Scores:
Pred. No.: 20.5 Length: 186
Score: 53.00 Matches: 15
Percent Similarity: 34.43% Conservative: 6
Best Local Similarity: 24.59% Mismatches: 16
Query Match: 16.21% Indels: 24
DB: 1 Gaps: 3

US-09-939-293-1_COPY_56_239 (1-184) x US-08-089-458B-6 (1-186)

QY 33 TGTTCGTGTGTGGCTACTTATAGAGCGGTCTCTCAGATTGATTAAGCATGGCA 92
|||||
Db 114 CysSerCysIeu-----ProGlyTTrpPheCysAlaThrAspSerSerIysThr 129
QY 93 CAAACCTGTGACGATGTGGCTT----- 113
:::|::|
Db 130 GluAspCysArgAspCysIleProLysArgLysCysProCysGlyTyrPheGlyGlyIle 149
QY 114 -----TGGAGTAAACCTGTGTGC-----GGTTCTATTTGCACAGAAATC 152
|||||
Db 150 AspGluLeuGlnLysProLeuCysLysSerCysCysValGlyGluTyrCysAspAspIle 169
QY 153 AGA 155
|||
Db 170 Arg 170

Search completed: February 16, 2003, 03:13:11
Job time : 24.2836 secs

PF 06-JAN-2000; 2000US-0479309.
XX
PR 06-JAN-2000; 2000US-0479309.
XX
XX (TEXA) UNIV TEXAS SYSTEM.
PA
XX Wang X, Du C;
PI
XX WPI: 2000-586350/55.
DR N-PSDB; AAA94860.
XX
PT Novel caspase regulatory polypeptides useful for screening binding
agents specific for the polypeptides which are useful for diagnosis and
also for treating apoptosis associated diseases
PS Claim 1; column 23-24; 16pp; English.
XX
CC The present sequence is the human Smac protein. Its coding sequence
CC was isolated by purifying the protein and searching a HeLa cell cDNA
CC library for sequences which bound to probes based upon it. Smac is a
CC mitochondrial protein which is released into the cytosol during
CC apoptosis, and acts as a caspase-3 activator. The protein and its coding
CC sequence can be used to modulate the expression and function of caspases
CC and their activators, and also can be used as drug targets and regulators
CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
CC and neurodegenerative diseases.
XX
SQ Sequence 239 AA;

Alignment Scores:
Pred. No.: 4.68e-36 Length: 239
Score: 327.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-939-293-1_COPY_56_239 (1-184) x AAB26210 (1-239)
QY 1 ACTTCATCTTCAGGTACAGACAGTGTGTTGCTGTTGGCTAACTTTAAGAG 60
DB 13 Thrsrphrphrargtyrarglncysleucysvalprovalvalaalaasnphelysls 32
QY 61 CGGTGTTTCTCAGAAATTGATTAAGACCAGGCAAAAACGTGTACGATTGGCTTGGAGTA 120
DB 33 ArgcyspheserGluleuilearProtrphisthrValThrleGlypheclyVal 52
QY 121 ACCGTGTGGGGTCTTATTCACAGAAATGAGAGCCCTTCCCTTGTAGTAGAGA 180
DB 53 ThrleucysAlaValProlelaGlnlySerGluProHisserleuSerGlnAla 72
QY 181 TTG 183
DB 73 Leu 73

RESULT 2
AAU78447
ID AAU78447 standard; Protein; 239 AA.
XX
AC AAU78447;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; tumour.
XX
OS Homo sapiens.
XX
PN WO200216418-A2.

PD 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US26492.
PF
XX
XX 24-AUG-2000; 2000US-227735P.
PR
XX (UYJE-) UNIV JEFFERSON THOMAS.
PA
XX Alnemrl ES;
PI
XX WPI: 2002-304115/34.
DR N-PSDB; ABK15451.
XX
PT Novel Smac peptides and polynucleotides encoding the peptides, useful
PT for stimulating apoptosis in neoplastic or tumour cell which
PT overexpresses inhibitor of caspase, and for identifying apoptosis
PT modulating compounds
PS Claim 36; page 73-74; 78pp; English.
XX
CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac protein.
XX
SQ Sequence 239 AA;

Alignment Scores:
Pred. No.: 4.68e-36 Length: 239
Score: 327.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 23 Gaps: 0

US-09-939-293-1_COPY_56_239 (1-184) x AAU78447 (1-239)
QY 1 ACTTCATCTTCAGGTACAGACAGTGTGTTGCTGTTGGCTAACTTTAAGAG 60
DB 13 Thrsrphrphrargtyrarglncysleucysvalprovalvalaalaasnphelysls 32
QY 61 CGGTGTTTCTCAGAAATTGATTAAGACCAGGCAAAAACGTGTACGATTGGCTTGGAGTA 120
DB 33 ArgcyspheserGluleuilearProtrphisthrValThrleGlypheclyVal 52

QY 121 ACCCTGTCGGTTCCTATTGCACAGAAATCAGAGCCCTCATTCCTTACTAGTGAAGCA 180
 Db 53 ThleucysAlaValProIleAlaGlnLysSerGluProHisSerLeuSerSerGluAla 72
 OY 181 TTG 183
 |||
 Db 73 Leu 73

RESULT 3
 AAB54139
 ID AAB54139 standard; Protein; 227 AA.
 XX AAB54139;
 AC AAB54139;
 DT 09-MAR-2001 (first entry)
 DE Human pancreatic cancer antigen protein sequence SEQ ID NO:591.
 XX
 XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW detection; diagnosis; identification; cytostatic; neuroprotective;
 KW neoplastic; immunomodulatory; relaxant; contraceptive; gynaecological;
 KW antiinflammatory; cardiac; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative.
 XX
 OS Homo sapiens.
 XX
 PN WO200055320-A1.
 XX
 XX 21-SEP-2000.
 PD
 XX 08-MAR-2000; 2000WO-US05989.
 XX
 XX 12-MAR-1999; 99US-0124270.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI Rosen CA, Ruben SM;
 DR WPI; 2000-579444/54.
 DR N-PSDB; AAC98904.
 XX
 XX New nucleic acid that is a pancreatic cancer antigen for preventing,
 PT treating, or ameliorating a medical condition, particular pancreatic
 PT cancer, or for use in assays for diagnosing a pathological condition -
 XX
 PS Claim 11; Page 1027-1028; 1379pp; English.

AAC98773 to AAC99231 encode the human pancreatic cancer associated
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,
 CC neuroprotective, neoplastic, immunomodulatory, relaxant, contraceptive,
 CC gynaecological, cardiac and antiinflammatory activities, and can be used
 CC in gene therapy. The polynucleotide and proteins can be used for
 CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in chromosome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect and target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.
 XX

SO Sequence 227 AA;
 Alignment Scores:
 Pred. No.: 2,24e-35 Length: 227
 Score: 322.00 Matches: 60
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.47% Indels: 0
 DB: 21 Gaps: 0

US-09-939-293-1_COPY_56_239 (1-184) x AAB54139 (1-227)

QY 4 TCATTCTTCAGGTACACAGAGTGTTCCTGCTTCCTTGGCTTAACCTTAAGACCG 63
 |||
 Db 2 SerpHeaRgYrYrARgGlnCysLeucysValProValAlaLaanPhelYsYsArg 21
 |||
 OY 64 TGTTCACGAATGATAGACCATGGACACAAACCTGACATGGCTTGGAGTAAAC 123
 |||
 Db 22 CysPheSerGluLeuIleArgProTrpHisLysThrValThrIleGlyPheGlyValThr 41
 |||
 OY 124 CTGTGTCGGTTCCTATTGCACAGAAATCAGAGCCCTCATTCCTTACTAGTGAAGCAT 183
 |||
 Db 42 LeucysAlaValProIleAlaGlnLysSerGluProHisSerLeuSerSerGluAlaLeu 61
 |||

RESULT 4
 ABB76208
 ID ABB76208 standard; Peptide; 20 AA.
 XX
 XX ABB76208;
 AC ABB76208;
 XX
 DT 09-AUG-2002 (first entry)
 DE Human smac (DIABLO) derived peptide.
 XX
 XX DIABLO; smac; inhibitor of apoptosis protein; IAP; apoptosis;
 KW human; cancer; cytostatic.
 KW
 OS Homo sapiens.
 XX
 OS
 FH Key location/Qualifiers
 FT Modified-site 20
 FT /note="optional C-terminal protecting group"
 XX
 XX WO200230959-A2.
 XX
 XX 18-APR-2002.
 PD
 XX 12-OCT-2001; 2001WO-US32121.
 PF
 XX 13-OCT-2000; 2000US-0687549.
 PR
 XX (ABBO) ABBOTT LAB.
 PA
 PI Fesik SW, Meadows RP, Betz SP, Liu Z, Olejniczak EF, Sun C;
 DR WPI; 2002-444169/47.
 XX
 XX Novel peptide derived from wild-type human second mitochondria derived
 PT activator of caspase protein useful for identifying candidate
 PT substances to kill cancerous cells -
 XX
 PS Claim 5; Page 7; 26pp; English.

The present sequence is a peptide derived from wild-type human
 CC second mitochondria derived activator of caspase (smac), also known
 CC as direct inhibitor of apoptosis binding protein with low PI
 CC (DIABLO). The peptide is one of 12 claimed smac (DIABLO)-derived
 CC peptides (see ABB76208-19) which bind to the Bir2 and Bir3 domain
 CC of XIAP, an inhibitor of apoptosis protein (IAP) family member.
 CC Kd values for Bir-3 and Bir-2 are 0.69 +/- 0.05 uM and 6.7 +/- 0.7
 CC uM, respectively, for the present peptide, compared with 0.42 +/-
 CC 0.02 uM and 2.3 +/- 0.3 uM, respectively, for full-length smac.
 CC Modification of the N-terminal alanine destroys binding affinity to

CC XIAP. For example, N-terminal acetylation of the present peptide,
CC replacement of the N-terminal alanine with glycine, propionic acid
CC or isobutyric acid all resulted in Kd values for Bir-3 and for Bir-2
CC of over 1,000 nM. The claimed peptides can be used to identify
CC candidate substances which induce or promote apoptosis in cells.
CC The assay involves determination of the ability of candidate
CC compounds to disrupt the binding interaction between a smac (DIABLO)
CC peptide and an IAP family member.
XX
SQ Sequence 20 AA:

Alignment Scores:
Pred. No.: 0.00298 Length: 20
Score: 86.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.30% Indels: 0
DB: 23 Gaps: 0

US-09-939-293-1_COPY_56_239 (1-184) x ABB76208 (1-20)
OY 130 GCGGTCCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGAGCATG 183
Db 1 AlaValProIleAlaGlnIlysserGluProHisSerLeuSerSerLualaleu 18

RESULT 5
AAU78435
ID AAU78435 standard; Peptide; 30 AA.
XX
AC AAU78435;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac, mutant Smac-N30.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; mutant; tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2002-304115/34.
XX
PS
XX
XX The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the

CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac mutant
CC Smac-N30.
XX
SQ Sequence 30 AA:

Alignment Scores:
Pred. No.: 0.00323 Length: 30
Score: 86.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.30% Indels: 0
DB: 23 Gaps: 0

US-09-939-293-1_COPY_56_239 (1-184) x AAU78435 (1-30)
OY 130 GCGGTCCTATTGCACAGAAATCAGAGCCTCATTCCTTAGTAGAGCATG 183
Db 1 AlaValProIleAlaGlnIlysserGluProHisSerLeuSerSerLualaleu 18

RESULT 6
AAU78439
ID AAU78439 standard; Peptide; 35 AA.
XX
AC AAU78439;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac, peptide Smac-N35.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2002-304115/34.
XX
PS
XX
XX Novel Smac peptides and polynucleotides encoding the peptides, useful
PT for stimulating apoptosis in neoplastic or tumour cell which

PT overexpresses inhibitor of caspase, and for identifying apoptosis
modulating compounds -
XX
PS Example 4; Page 47; 78pp; English.
XX
CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDFYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac peptide
CC Smac-N35.
XX
SQ Sequence 35 AA:
XX
Alignment Scores:
Pred. No.: 0.00333 Length: 35
Score: 86.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.30% Indels: 0
DB: Gaps: 0
US-09-939-293-1_COPY_56_239 (1-184) x AAU78436 (1-35)
OY 130 GCGGTCCTATTGCACAGAAATCAGAGCTCATTCCTTAGTAGTGAAGCATTTG 183
DB 1 AlavaiProilealagInlysserGIupronHisSerleuserSerGIualaleu 18
RESULT 7
AAU78436
ID AAU78436 standard; Peptide; 39 AA.
XX
AC AAU78436;
XX
DT 18-JUN-2002 (first entry)
XX
DE Inhibitor of apoptosis (IAP) protein Smac, mutant Smac-N35.
XX
KW Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
KW neoplastic cell; mutant; tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX

PN WO200216418-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26492.
XX
PR 24-AUG-2000; 2000US-227735P.
XX
PA (UYE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemr1 ES;
XX
DR WPI; 2002-304115/34.
XX
PT Novel Smac peptides and polynucleotides encoding the peptides, useful
PT for stimulating apoptosis in neoplastic or tumour cell which
PT overexpresses inhibitor of caspase, and for identifying apoptosis
modulating compounds -
XX
PS Example 3; Fig 7; 78pp; English.
XX
CC The invention relates to an isolated Smac peptide or polypeptide (I)
CC and an isolated nucleic acid (II) encoding (I). Also described is a
CC method of identifying a compound that inhibits apoptosis, comprising:
CC (a) separately contacting several cell populations expressing a
CC cytosolic Smac (a Smac isoform that begins with MKSDFYF sequence,
CC replacing the mitochondrial targeting sequence (residues 1-55 of (I)),
CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
CC domain) with a compound to be tested for apoptotic inhibiting activity;
CC (b) incubating the cell populations with a direct stimulus of the cell
CC death pathway; and (c) measuring the specific apoptotic activity of the
CC cell populations, where inhibition of the specific apoptotic activity is
CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
CC tumour cell which overexpresses an inhibitor of caspase, where the
CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
CC mediated apoptosis which involves contacting a cell transformed or
CC transfected with a vector expressing (I) with a candidate inhibitor or
CC candidate enhancer; and detecting cell viability, where an increase in
CC cell viability indicates the presence of an inhibitor and a decrease in
CC cell viability indicates the presence of an enhancer. Optionally, the
CC method involves detecting the presence of large and small caspase
CC subunits after contacting cell transformed with the vector expressing
CC (I), with the candidate compound. A decrease in processing indicates the
CC presence of an inhibitor and an increase in the processing indicates the
CC presence of an enhancer. Preferably, the large and small subunits of
CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
CC identifying a compound that inhibits Smac binding to Smac-binding
CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
CC or a full-length IAP). (II) is useful in gene therapy techniques. The
CC present sequence represents the amino acid sequence of Smac mutant
CC Smac-N39.
XX
SQ Sequence 39 AA:
XX
Alignment Scores:
Pred. No.: 0.00341 Length: 39
Score: 86.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.30% Indels: 0
DB: Gaps: 0
US-09-939-293-1_COPY_56_239 (1-184) x AAU78436 (1-39)
OY 130 GCGGTCCTATTGCACAGAAATCAGAGCTCATTCCTTAGTAGTGAAGCATTTG 183
DB 1 AlavaiProilealagInlysserGIupronHisSerleuserSerGIualaleu 18
RESULT 8

AAU78430
 ID AAU78430 standard; Peptide; 40 AA.
 AC AAU78430;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 XX Inhibitor of apoptosis (IAP) protein Smac, N-terminal peptide.
 DE
 XX Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
 KM Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
 KM neoplastic cell; tumour.
 XX
 XX Homo sapiens.
 OS
 XX MO200216418-A2.
 PN
 XX 28-FEB-2002.
 PD
 XX 24-AUG-2001; 2001WO-US26492.
 PF
 XX 24-AUG-2000; 2000US-227735P.
 PR
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 PA
 XX AInesmri ES;
 PI
 XX WPI; 2002-304115/34.
 DR
 XX
 XX Novel Smac peptides and polynucleotides encoding the peptides, useful
 PT for stimulating apoptosis in neoplastic or tumour cell which
 PT overexpresses inhibitor of caspase, and for identifying apoptosis
 PT modulating compounds -
 XX
 XX Example 3; Fig 7; 78pp; English.
 PS
 XX The invention relates to an isolated Smac peptide or polypeptide (I)
 CC and an isolated nucleic acid (II) encoding (I). Also described is a
 CC method of identifying a compound that inhibits apoptosis, comprising:
 CC (a) separately contacting several cell populations expressing a
 CC cytosolic Smac (a Smac isoform that begins with MKSDYF sequence,
 CC replacing the mitochondrial targeting sequence (residues 1-35 of (I)),
 CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
 CC domain) with a compound to be tested for apoptotic inhibiting activity;
 CC (b) incubating the cell populations with a direct stimulus of the cell
 CC death pathway; and (c) measuring the specific apoptotic activity of the
 CC cell populations, where inhibition of the specific apoptotic activity is
 CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
 CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
 CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
 CC tumour cell which overexpresses an inhibitor of caspase, where the
 CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
 CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
 CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
 CC mediated apoptosis which involves contacting a cell transformed or
 CC transfected with a vector expressing (I) with a candidate inhibitor or
 CC candidate enhancer; and detecting cell viability, where an increase in
 CC cell viability indicates the presence of an inhibitor and a decrease in
 CC cell viability indicates the presence of an enhancer. Optionally, the
 CC method involves detecting the presence of large and small caspase
 CC subunits after contacting cell transformed with the vector expressing
 CC (I), with the candidate compound. A decrease in processing indicates the
 CC presence of an inhibitor and an increase in processing indicates the
 CC presence of an enhancer. Preferably, the large and small subunits of
 CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
 CC identifying a compound that inhibits Smac binding to Smac-binding
 CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
 CC or a full-length IAP). (II) is useful in gene therapy techniques. The
 CC present sequence represents the N-terminal amino acid sequence of Smac
 CC protein.
 XX
 SQ Sequence 40 AA;

Alignment Scores:
 Pred. No.: 0.00342 Length: 40
 Score: 86.00 Matches: 18
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 26.30% Indels: 0
 DB: 23 Gaps: 0
 US-09-939-293-1_COPY_56_239 (1-184) x AAU78430 (1-40)
 QY 130 GCGGTCCTCATTCGACAGAAATCAGAGCCTCATTCCTTAGTGAACATTG 183
 DB 1 AAlavAlProIIeAlaGInlySserGIuProHISserLeuSerSerCIuAlaLeu 18
 RESULT 9
 AAB13251
 ID AAB13251 standard; Protein; 197 AA.
 AC AAB13251;
 XX
 XX 12-JAN-2001 (first entry)
 DT
 XX Human metabolic enzyme sequence #1.
 DE
 XX Human; metabolic enzyme; AKT kinase; daf-18; insulin signalling pathway;
 KM daf-2; age-1; insulin receptor; PI 3-kinase; PKB kinase;
 KM PTEN lipid phosphatase; antidiabetic; anorectic; obesity; diabetes.
 XX
 XX Homo sapiens.
 OS
 XX WO200033068-A1.
 PN
 XX 08-JUN-2000.
 PD
 XX 02-DEC-1999; 99WO-US28529.
 PF
 XX 03-DEC-1998; 98US-0205658.
 PR
 XX (GEHO) GEN HOSPITAL CORP.
 PA
 XX Ruvkun G, Ogg S;
 PI
 XX WPI; 2000-423022/36.
 DR
 XX
 XX Diagnosing and treating obesity and impaired glucose tolerance using
 PT modulators of daf-18 expression and/or activity -
 PT
 XX
 XX Disclosure: Page 168; 402pp; English.
 PS
 XX The present sequence is part of a key metabolic enzyme from *Ascoris suum*
 CC which shows homology to a protein from *Caenorhabditis elegans* genome. A
 CC number of C. elegans genes have been identified as homologues of genes in
 CC the mammalian insulin signalling pathway. The C. elegans age-1 gene
 CC encodes a homologue of the mammalian PI 3-kinase whilst daf-2 encodes a
 CC homologue of the mammalian insulin receptor. The C. elegans AKT kinase
 CC and PKB kinase act downstream of daf-2 and age-1, just as their mammalian
 CC homologues act downstream of insulin signalling. The C. elegans PTEN
 CC lipid phosphatase homologue, DAF-18, has been found to act upstream
 CC of AKT in the pathway. This discovery has enabled mammalian PTEN action
 CC to be mapped to the insulin signalling pathway. Conserved DAF motifs can
 CC be used to design probes to identify mammalian DAF homologues and thus to
 CC identify individuals with a predisposition toward the development of
 CC glucose intolerance conditions, such as obesity and diabetes.
 XX
 SQ Sequence 197 AA;
 Alignment Scores:
 Pred. No.: 3.02 Length: 197
 Score: 65.50 Matches: 15
 Percent Similarity: 51.02% Conservative: 10
 Best Local Similarity: 30.61% Mismatches: 13
 Query Match: 20.03% Indels: 1
 DB: 21 Gaps: 1

```

US-09-939-293-1_COPY_56_239 (1-184) x AAB13251 (1-197)
OY 67 TTCTCAGAAATTGTAAGACCATGCGCACAAACTGTGACATTGGC----- 111
DB 9 PHeSerGluIleMetGlnProTrpAlaGlnThrValValaIcIYArGaIaArgLeuGly 28
OY 112 -----TTTGAGTAACCCGTGTGCGGTCCCTATTGTCACAGAAATCA 153
DB 29 GYLIIeProValGIyValaValaIcIuThrArgThrValaIcIuSerValaIProAla 48
OY 154 GAGCCTCATTCCTTAGTAGTAGCAACA 180
DB 49 AsProAlaAsnLeuAspSerGluAla 57

RESULT 10
AAB86033
ID AAB86033 standard; Protein; 2288 AA.
AC AAB86033;
XX
XX 13-JUL-2001 (first entry)
DE Bovine acetyl-coenzyme A carboxylase-alpha protein fragment.
XX
XX Acetyl-coenzyme A carboxylase-alpha; acetyl-CoA carboxylase alpha;
XX bovine; milk gland-specific promoter; Accalpa; milk production; sheep;
XX goat; fat content; genotyping; lactation.
XX
XX Bos taurus.
XX
XX DE19946173-A1.
XX
XX 05-APR-2001.
XX
XX 20-SEP-1999; 99DE-1046173.
XX
XX 20-SEP-1999; 99DE-1046173.
XX
XX (BIOL-) FORSCHUNGSINSTITUT BIOLOGIE LANDMIRTSCHA.
XX
XX Seyfert HM;
XX
XX WPI: 2001-258968/27.
XX
XX N-PSDB: AAF88002.
XX
XX New nucleic acid encoding bovine acetyl coenzyme A carboxylase alpha
XX and its promoter, for milk-specific production of proteins and for
XX regulating fat content of milk
XX
XX Claim 2c: Page 24-30; 44pp; German.
XX
XX This invention describes a novel milk gland-specific promoter of the
XX bovine acetyl coenzyme A carboxylase alpha gene (Accalpa), or its
XX fragments, which are used to control expression of foreign genes. When
XX the promoter (or the Accalpa structural gene) is replaced, at least in
XX part, by a sequence that is altered by deletion or substitution, then
XX expression of Accalpa in the milk gland is reduced and the milk produced
XX (by cattle, sheep or goats) has reduced fat content. Analysis of the
XX polymorphic 933-966 region of (Accalpa) is useful for genotyping
XX animals, producing a genotype that is directly correlated with Accalpa
XX expression during lactation and with fat content of the milk. This
XX makes it possible to generate populations of cattle that produce milk of
XX high or low fat content by classical breeding methods. Accalpa is a
XX lactation-specific, inducible promoter for expressing foreign proteins
XX in the milk and, when modified, results in milk of reduced fat content,
XX which facilitates recovery of proteins. This sequence represents a
XX fragment from bovine Accalpa which contains a fragment of exon 5A
XX exon 6 and exon 7.
XX
XX Sequence 2288 AA;
XX

```

```

Pred. No.: 6.73 Length: 2288
Score: 64.50 Matches: 15
Percent Similarity: 51.02% Conservative: 10
Best Local Similarity: 30.61% Mismatches: 13
Query Match: 19.72% Indels: 11
DB: 22 Gaps: 1

US-09-939-293-1_COPY_56_239 (1-184) x AAB6033 (1-2288)
OY 67 TTCTCAGAAATTGTAAGACCATGCGCACAAACTGTGACATTGGC----- 111
DB 1902 PHeSerGluIleMetGlnProTrpAlaGlnThrValValaIcIYArGaIaArgLeuGly 1921
OY 112 -----TTTGAGTAACCCGTGTGCGGTCCCTATTGTCACAGAAATCA 153
DB 1922 GYLIIeProValGIyValaValaIcIuThrArgThrValaIcIuSerIleProAla 1941
OY 154 GAGCCTCATTCCTTAGTAGTAGCAACA 180
DB 1942 AsProAlaAsnLeuAspSerGluAla 1950

RESULT 11
AAB92922
ID AAB92922 standard; Protein; 186 AA.
XX
XX AAB92922;
AC AAB92922;
XX
XX 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:11570.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.
XX
XX primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs.
XX
XX Claim 8: SEQ ID 11570; 2537bp + CD ROW; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in

```

Alignment Scores:

| | |
|--|---|
| CC | the specification. The primer sets can be used in antisense therapy and |
| CC | in gene therapy. The primers are useful for synthesising polynucleotides |
| CC | particularly full-length cDNAs. The primers are also useful for the |
| CC | detection and/or diagnosis of the abnormality of the proteins encoded by |
| CC | the full-length cDNAs. The primers allow obtaining of the full-length |
| CC | cDNAs easily without any specialised methods. AAH03166 to AAH13628 and |
| CC | AAH13633 to AAH18742 represent human CDNA sequences; AAB92446 to |
| CC | AAB95693 represent human amino acid sequences; and AAH13629 to AAH13632 |
| CC | represent oligonucleotides, all of which are used in the exemplification |
| CC | of the present invention. |
| XX | |
| SO | Sequence 186 AA; |
| XX | |
| Alignment Scores: | |
| Pred. NO.: | 6.57 |
| Score: | 63.00 |
| Percent Similarity: | 100.00% |
| Best Local Similarity: | 100.00% |
| Query Match: | 19.27% |
| DB: | 22 |
| | Gaps: 0 |
| US-09-939-293-1-COPY_56-239 (1-184) x AAB92922 (1-186) | |
| QY | 145 CAGAAATCAGAGCCTCATTCCTTAGTAGAGCATTTG 183 |
| DB | |
| | 8 GlnlySerGluProHisSerLeuSerGluAlaIeu 20 |
| RESULT 12 | |
| ID | AAB938293 |
| AC | AAB938293; standard; Protein; 190 AA. |
| XX | |
| DT | 24-JUL-2002 (first entry) |
| XX | |
| DE | Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3138. |
| XX | |
| KW | Staphylococcus epidermidis; open reading frame; ORF; bacterial infection. |
| KW | antibacterial; gene therapy. |
| OS | Staphylococcus epidermidis. |
| XX | |
| PN | US6380370-B1. |
| XX | |
| PD | 30-APR-2002. |
| XX | |
| PF | 13-AUG-1998; 98US-0134001. |
| XX | |
| PR | 14-AUG-1997; 97US-055779P. |
| PR | 08-NOV-1997; 97US-064964P. |
| XX | |
| PA | (GENO-) GENOME THERAPEUTICS CORP. |
| XX | |
| PI | Doucette-Stamm LA, Bush D; |
| XX | |
| DR | WPI: 2002-381255/41. |
| DR | N-PSDB: AAB90838. |
| XX | |
| PT | Novel isolated nucleic acid encoding a Staphylococcus epidermidis |
| XX | polypeptide, useful for diagnosing and treating bacterial infections - |
| XX | |
| PS | Disclosure; SEQ ID 3138; 267pp; English. |
| XX | |
| CC | AAB90538 to AAB93374 represent Staphylococcus epidermidis open reading |
| CC | frame (ORF) nucleic acid sequences which encode the amino acid sequences |
| CC | given in AAB95124 to AAB97960. The S. epidermidis sequences have |
| CC | antibacterial activity and can be used in gene therapy. The sequences |
| CC | can also be used in the diagnosis and treatment of bacterial infections, |
| CC | particularly S. epidermidis infections. The sequences can be used to |
| CC | screen for compounds able to interfere with the S. epidermidis life |
| CC | cycle or inhibit S. epidermidis infection. |
| CC | N.B. The sequence data for this patent did not form part of the printed |
| CC | specification, but was obtained in electronic format directly from the |

```

CC  USPTO web site.
XX
SQ  Sequence 190 AA;

Alignment Scores:
Pred. No.: 9.05 Length: 190
Score: 62.00 Matches: 12
Percent Similarity: 72.00% Conservative: 6
Best Local Similarity: 48.00% Mismatches: 7
Query Match: 18.96% Indels: 0
DB: 23 Gaps: 0

US-09-939-293-1_COPY_56_239 (1-184) x ABP38293 (1-190)

OY 10 TTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAACTTAAAGAGCGGTGTTTC 69
    ::::||||| ||| :::::||||| ||||| ||||| |||||
Db 2 TtgcIntYrAspleuAaBleuLeuVaiYsIleuLysLysArgcysPhe 21
OY 70 TCAGATTGATAGA 84
    ||:::|||||
Db 22 SerLysMetIleArg 26

RESULT 13
ABB59490
ID ABB59490 standard; protein; 2348 AA.
XX
AC ABB59490;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SRQ ID NO 5262.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
MO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EM;
XX
DR WI: 2001-656860/75.
XX
DR N-PSDB; ABL03593.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PT Interactions -
XX
Disclousre; SEQ ID NO 5262; 21pp + Sequence Listing; English.
XX
PS
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2348 AA;

```


CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate hematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX

SQ Sequence 2486 AA:

Alignment Scores:

| | | | |
|------------------------|--------|---------------|------|
| Pred. No.: | 17.6 | Length: | 2486 |
| Score: | 61.50 | Matches: | 16 |
| Percent Similarity: | 42.59% | Conservative: | 7 |
| Best Local Similarity: | 29.63% | Mismatches: | 20 |
| Query Match: | 18.81% | Indels: | 11 |
| DB: | 22 | Gaps: | 1 |

US-09-939-293-1_COPY_56_239 (1-184) x AAU32848 (1-2486)

```

QY 52 TTTAAGACGGGTGTTCTTCAGATTGATTAAGACCATGGCACAAACTGTGACGATTGGC 111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2094 PHEASPHISGLYSERPHGLSGIULEMETALAIPROTIRPALAGLNRHRYVALYALThrgly 2113
QY 112 -----TTTGAGTAACCCCTGTGCGGTTCCCT 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2114 ARGALARGLEUGLYGLYILEPROVALIGLYVALIIEALAVAILGUTHRARGHRYVALIGLU 2133
QY 139 AFTGCACAGAAATCAGAGCCCTCATTCCTTAAGTAGTGAAGCA 180
    ::::|::| ::::|::| ::::|::| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2134 VALAIAVALPROALASPProAlaAsnLeuAspSerGluAla 2147

```

Search completed: February 16, 2003, 03:06:28
 Job time : 63.791 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 05:05:07 : Search time 3213 Seconds

(without alignments)
1666.641 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239
Perfect score: 184
Sequence: 1 actcattctctcaggtacag.....ccttagtaggaagcattga 184

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 1169814

Minimum DB seq length: 0
Maximum DB seq length: 184

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: GenEmbl:*
2: gb_ba:*
3: gb_htg:*
4: gb_in:*
5: gb_om:*
6: gb_ov:*
7: gb_pat:*
8: gb_ph:*
9: gb_pl:*
10: gb_pr:*
11: gb_ro:*
12: gb_sts:*
13: gb_sy:*
14: gb_un:*
15: gb_vi:*
16: em_ba:*
17: em_fun:*
18: em_hum:*
19: em_in:*
20: em_mu:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_un:*
30: em_vi:*
31: em_htg_hum:*
32: em_htg_inv:*
33: em_htg_other:*
34: em_htg_mus:*
35: em_htg_pln:*
36: em_htg_rnd:*
37: em_htg_mam:*
38: em_htg_vrt:*
39: em_sy:*
40: em_htgo_hum:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 29 | 15.8 | 162 | 11 | AF225190 Sus scrofa |
| 2 | 25.4 | 13.8 | 120 | 4 | AF033020 Sus scrofa |
| 3 | 25 | 13.6 | 163 | 11 | G44212 WtAF-3856-S |
| 4 | 23.8 | 12.9 | 114 | 6 | AX438757 Sequence |
| 5 | 23.6 | 12.8 | 65 | 6 | AX485057 Sequence |
| 6 | 23.6 | 12.8 | 138 | 9 | HUMTCAAR |
| 7 | 23.6 | 12.8 | 178 | 11 | HUMCS122 |
| 8 | 23.4 | 12.7 | 149 | 11 | AU028588 |
| 9 | 23.4 | 12.7 | 175 | 11 | G19018 |
| 10 | 23.2 | 12.6 | 174 | 6 | I51849 |
| 11 | 23 | 12.5 | 156 | 11 | AF275557 |
| 12 | 22.8 | 12.4 | 90 | 9 | D50749 |
| 13 | 22.8 | 12.4 | 112 | 9 | AF088775 |
| 14 | 22.8 | 12.4 | 112 | 9 | AF088796 |
| 15 | 22.8 | 12.4 | 121 | 9 | AF088715 |
| 16 | 22.8 | 12.4 | 123 | 9 | AF088689 |
| 17 | 22.8 | 12.4 | 134 | 9 | AF088462 |
| 18 | 22.8 | 12.4 | 136 | 9 | AF088372 |
| 19 | 22.8 | 12.4 | 137 | 9 | AF088385 |
| 20 | 22.8 | 12.4 | 138 | 9 | AF088370 |
| 21 | 22.8 | 12.4 | 138 | 9 | HS1AB2710 |
| 22 | 22.8 | 12.4 | 144 | 9 | AF088320 |
| 23 | 22.8 | 12.4 | 145 | 9 | AF088631 |
| 24 | 22.8 | 12.4 | 148 | 9 | AF088471 |
| 25 | 22.8 | 12.4 | 151 | 9 | AF088471 |
| 26 | 22.8 | 12.4 | 182 | 9 | HS275941 |
| 27 | 22.6 | 12.3 | 107 | 3 | AF226307 |
| 28 | 22.6 | 12.3 | 111 | 9 | HSU39099 |
| 29 | 22.6 | 12.3 | 133 | 11 | AU025617 |
| 30 | 22.6 | 12.3 | 136 | 11 | AU029131 |
| 31 | 22.6 | 12.3 | 137 | 5 | AF447866 |
| 32 | 22.4 | 12.2 | 115 | 9 | AF088717 |
| 33 | 22.4 | 12.2 | 123 | 11 | HS144YTC11 |
| 34 | 22.4 | 12.2 | 147 | 4 | AF389372 |
| 35 | 22.4 | 12.2 | 158 | 5 | AY052425 |
| 36 | 22.4 | 12.2 | 158 | 5 | AY052426 |
| 37 | 22.2 | 12.1 | 159 | 1 | BACAPAP |
| 38 | 22.2 | 12.1 | 169 | 6 | AX185907 |
| 39 | 22.2 | 12.1 | 177 | 11 | G18547 |
| 40 | 22 | 12.0 | 140 | 11 | AU048140 |
| 41 | 22 | 12.0 | 150 | 3 | LAU72034 |
| 42 | 22 | 12.0 | 167 | 3 | SUURS090 |
| 43 | 22 | 12.0 | 182 | 11 | BOVISTST02 |
| 44 | 21.8 | 11.8 | 138 | 10 | RRTRO4A21 |
| 45 | 21.8 | 11.8 | 149 | 11 | AF253755 |

ALIGNMENTS

RESULT 1
AF225190
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AF225190 162 bp DNA linear STS 25-APR-2000
Sus scrofa chromosome 3 clone SW2618 map 3p1.2-p1.1, sequence
tagged site.
AF225190
AF225190.1 GI:7643965
Sus scrofa.
Sus scrofa.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 162)
Alexander,L.J., Troyer,D.L., Rohrer,G.A., Smith,T.P., Schook,L.B.
and Beattie,C.W.

TITLE Physical assignments of 68 porcine cosmid and lambda clones containing polymorphic microsatellites
JOURNAL Mamm. Genome 7 (5), 368-372 (1996)
MEDLINE 96269605
PUBMED 8661726
REFERENCE 2 (bases 1 to 162)
AUTHORS Alexander,L.J., Troyer,D.L., Rohrer,G.A., Smith,T.P.L., Schook,L.B. and Beatlie,C.W.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2000) USDA, ARS, US Meat Animal Research Center, P. O. Box 166, Clay Center, NE 68933-0166, USA
FEATURES
source
1. 162
/organism="Sus scrofa"
/db_xref="taxon:9823"
/chromosome="3"
/map="3p1.2-p1.1"
/clone="SW2618"
primer_bind 3. 27
primer_bind complement(105..129)
BASE COUNT 25 a 11 c 55 g 71 t
ORIGIN
Query Match 15.8%; Score 29; DB 11; Length 162;
Best Local Similarity 63.8%; Pred. No. 1.4e+02;
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 24 GTGTTGTGTCCTGCTGTGCTACTTTAAGACGGGCTTTCACAGATTGATGAAAG 83
Db 65 GTGCTGTGTCGTGCTGTGAATCTGTTTGAAGAAAGTTGTCAGAGCTTTGAG 124
QY 84 AACATGCGA 92
Db 125 AACATGCGA 133
RESULT 2
AF033020/c 120 bp mRNA linear MAM 27-NOV-1997
LOCUS Sus scrofa cadherin-11 mRNA, partial cds.
DEFINITION AF033020
ACCESSION AF033020
VERSION AF033020.1 GI:2645716
KEYWORDS
SOURCE
ORGANISM
Sus scrofa.
Sus scrofa
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 120)
AUTHORS Lutz,D.A. and Zheng,J.J.
TITLE Expression of multiple cadherins in adult retinal pigment epithelial (RPE) cells
JOURNAL unpublished
REFERENCE 2 (bases 1 to 120)
AUTHORS Lutz,D.A. and Zheng,J.J.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-1997) Ophthalmology & Visual Sciences, University of Louisville, 301 E. Muhammad Ali Blvd., Louisville, KY 40202-1511, USA
FEATURES
source
Location/Qualifiers
1. 120
/organism="Sus scrofa"
/db_xref="taxon:9823"
/cell_type="pigment epithelium"
/tissue_type="retina"
/function="cell adhesion molecule"
/product="cadherin-11"
/codon_start=1
/protein_id="AAB87475.1"
/db_xref="GI:2645717"
/translation="SIQIYIEGRGSVAGSLSSLESATTTSDLDYDLQNWGPR"
BASE COUNT 28 a 33 c 33 g 26 t
ORIGIN

Query Match 13.8%; Score 25.4; DB 4; Length 120;
Best Local Similarity 61.2%; Pred. No. 2.2e+03;
Matches 41; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 7 TTCTCAGTAGACAGACGTGTTGTGTCCTGTTGCTACTTTAAGACGGGTGT 66
Db 107 TTCTCAGTAGACATGATGTCACAGTCTGAATCTGTGTCGCCACCTAAGAGACTCAG 48
QY 67 TTCACAG 73
Db 47 GACCCAG 41
RESULT 3
G44212/c 163 bp DNA linear STS 28-JAN-1999
LOCUS WIAF-3856-STS Human T Hudson SANGER Homo sapiens STS genomic, sequence tagged site.
DEFINITION G44212
ACCESSION G44212
VERSION G44212.1 GI:4193129
KEYWORDS STS.
SOURCE
ORGANISM
Homo sapiens.
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 163)
AUTHORS Wang,D.G., Fan,J.B., Siao,C.J., Berno,A., Young,P., Sapolsky,R., Ghandour,G., Perkins,N., Winchester,E., Spencer,J., Kruglyak,L., Stein,L., Hsie,L., Topaloglou,T., Hubbell,E., Robinson,E., Miltman,M., Morris,M.S., Shen,N., Kilburn,D., Rioux,J., Nusbam,C., Rozen,S., Hudson,T.J., Lipschutz,R., Chee,M. and Lander,E.S.
TITLE Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome
JOURNAL Science 280 (5366), 1077-1082 (1998)
MEDLINE 98248615
PUBMED 9582121
COMMENT
Synonyms: stsG26049
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu
Primer A: AACAGCTGACAGTCACG
Primer B: TCCTCAAGCCCAAGTCAT
STS size: 163
PCR Profile:
Presoak: 94 degrees C for 4.00 minutes
Denaturation: 94 degrees C for 50.0 seconds
Annealing: 58 degrees C for 1.50 minutes
Polymerization: 72 degrees C for 1.00 minutes
PCR Cycles: 30
Thermal Cycler: custom built by IAS, Costar, Cambridge MA
Protocol:
Template: 10 ng
Primer: each 5 pm
dNTPs: 4 mM
Tag Polymerase: 0.5 U
Total Vol: 20 uL
Buffer:
Mg2+: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
Gelatin: .001 %
location/Qualifiers
1. 163
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="62.50 cR from top of Chr11 linkage group"

```

/clone_lib="Human THudson SANGER"
/notes="human STS created from EST in the Sanger database"
STS
  primer_bind 1.163
  primer_bind 1.19
  primer_bind complement(144..163)
BASE COUNT 44 a 37 c 30 g 51 t 1 others
ORIGIN

Query Match 13.6%; Score 25; DB 11; Length 163;
Best Local Similarity 53.1%; Pred. No. 2.9e+03;
Matches 52; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 38 CTGTGGCTAACTTAAAGACGGTGTTCAGATGATGATGACGACACAAA 97
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 CTCCTTGGGTAACTGACGAGAGATGATACACAGAGAAATGTCATCTACTA 39
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 98 CTGTGACGATGGCTTTGGAGTAAACCTGTGTGGCGTT 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 GTGTGAATTTGTAAAGTGGCGTGTCTGTGACGCTGTT 1

RESULT 4
AX438757/c AX438757 114 bp DNA linear PAT 28-JUN-2002
LOCUS
DEFINITION Sequence 7172 from Patent WO0229113.
ACCESSION AX438757
VERSION AX438757.1 GI:21663565
KEYWORDS
SOURCE Bacillus clausii.
ORGANISM Bacillus clausii.
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Berke,R. and Clausen,I.G.
TITLES Methods for monitoring multiple gene expression
JOURNAL Patent: WO 0229113-A 7172 11-APR-2002;
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)
FEATURES
  source 1..114
    /organism="Bacillus clausii"
    /db_xref="taxon:79880" 29 t
BASE COUNT 34 a 25 c 26 g 29 t
ORIGIN

Query Match 12.9%; Score 23.8; DB 6; Length 114;
Best Local Similarity 57.3%; Pred. No. 7.3e+03;
Matches 43; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 85 CCATGGCACAACCTGTGACGATGGCTTGAAGTAACCTGTGTGGCGTTCTATTGCA 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 CCATGGCCCAATTAATGATGTTGCTGTGGCGGACGCTGATGATGACCATGGCG 37
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 145 CAGAAATCAGAGCCT 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 GTGAATGACAGCCT 22
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
AX485057 AX485057 65 bp DNA linear PAT 16-AUG-2002
LOCUS
DEFINITION Sequence 2357 from Patent WO02053728.
ACCESSION AX485057
VERSION AX485057.1 GI:22319341
KEYWORDS
SOURCE Candida albicans.
ORGANISM Candida albicans.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE 1
AUTHORS Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
TITLES Gene disruption methodologies for drug target discovery
JOURNAL Patent: WO 02053728-A 2357 11-JUL-2002;
Eliara Pharmaceuticals, Inc. (US)
FEATURES
  Location/Qualifiers

source 1..65
  /organism="Candida albicans"
  /db_xref="taxon:5476"
BASE COUNT 21 a 3 c 16 g 25 t
ORIGIN

Query Match 12.8%; Score 23.6; DB 6; Length 65;
Best Local Similarity 64.8%; Pred. No. 8.7e+03;
Matches 35; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 14 GGTACAGACAGTGTGTGTCTGTCTGTGGCTTAAGTAAAGCGGTGT 67
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 GGTAAAGACATTTGTGTGTGTGTATGATATCGAATGATATTAATTAATGAGAT 58
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
HUMTCRAAR/c HUMTCRAAR 138 bp mRNA linear PRI 14-DEC-2001
LOCUS
DEFINITION Homo sapiens T-cell receptor alpha (TCRA) mRNA (HLA-A3, 29; B7, 44;
DR 2..7), partial cds.
ACCESSION U34713
VERSION U34713.1 GI:1100137
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Homo sapiens.
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Barber,D.F., Lopez,D. and Lopez de Castro,J.A.
TITLES T cell receptor diversity in alloreactive responses against HLA-B27
(B*2705) is limited by multiple-level restrictions in both alpha
and beta chains
JOURNAL Eur.J. Immunol. 25 (9), 2479-2485 (1995)
MEDLINE 96011855
PUBMED 7589114
FEATURES
  source 1..138
    /organism="Homo sapiens"
    /isolate="SR"
    /db_xref="taxon:9606"
    /map="14q11.2"
    /haplotype="HLA-A3, 29; B7, 44; DR 2, 7"
    /cell_line="CTL 47569"
    /cell_type="T-cell, cytolytic"
    /tissue_type="blood"
    /dev_stage="adult"
    1..138
    /gene="TCRA"
    /gene="TCRA"
    <1..>138
    /gene="TCRA"
    /note="putative"
    /codon_start=1
    /product="T-cell receptor alpha"
    /protein_id="AA82661.1"
    /db_xref="GI:1100138"
    /db_xref="CDB:G00-120-404"
    /translation="RKSAADVISAQLGDSAMTFCAIMIGGSNYKLTFGKGLITVNP
    NT"
    1..69
    /gene="TCRA"
    /note="TCRA V delta 4 segment; G00-120-404; putative"
    70..133
    /gene="TCRA"
    /note="TCRA N + J alpha HAVT33 segment; G00-120-404;
    putative"
    134..138
    /gene="TCRA"
    /note="TCRA C alpha segment; G00-120-404; putative"
BASE COUNT 44 a 31 c 28 g 35 t
ORIGIN

Query Match 12.8%; Score 23.6; DB 9; Length 138;
Best Local Similarity 69.6%; Pred. No. 8.5e+03;

```

| | | | | | | | | | |
|--------------------------|---|---|-----|------------|-----------------|--------|----|------|----|
| Matches | 32: | Conservative | 0: | Mismatches | 14: | Indels | 0: | Gaps | 0: |
| OY | 139 | ATTGCACAGAAATCAGACCTTCATCCCTTAGTACTAGCAACTTGA | 184 | | | | | | |
| Db | 68 | ATTGCACAGAAATATCATCTTGTAGAGTCCCAAGTTGTGAAGCGAGA | 23 | | | | | | |
| RESULT 7 | | | | | | | | | |
| LOCUS | HUMC5122 | 178 bp | DNA | linear | STS 04-MAR-1997 | | | | |
| DEFINITION | Human chromosome 5 LANL STS 122, sequence tagged site. | | | | | | | | |
| ACCESSION | L28215 | | | | | | | | |
| VERSION | L28215.1 | GI:1160667 | | | | | | | |
| KEYWORDS | STS. | | | | | | | | |
| SOURCE | Homo sapiens. | | | | | | | | |
| ORGANISM | Homo sapiens. | | | | | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | | | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | | | | | |
| | 1 (bases 1 to 178) | | | | | | | | |
| | Grady,D.L., Robinson,D.L., Gersh,M., Nickerson,E., McPherson,J., | | | | | | | | |
| | Masmuth,J.J., Overhauser,J., Deaven,L.L. and Moyzis,R.K. | | | | | | | | |
| | The generation and regional localization of 303 new chromosome 5 | | | | | | | | |
| | sequence-tagged sites | | | | | | | | |
| | Genomics 32 (1), 91-96 (1996) | | | | | | | | |
| JOURNAL | Genomics 32 (1), 91-96 (1996) | | | | | | | | |
| MEDLINE | 96230329 | | | | | | | | |
| PUBMED | 8786125 | | | | | | | | |
| FEATURES | Location/Qualifiers | | | | | | | | |
| source | 1..178 | | | | | | | | |
| | /organism="Homo sapiens" | | | | | | | | |
| | /db_xref="taxon:9606" | | | | | | | | |
| | /chromosome="5" | | | | | | | | |
| | /tissue_lib="LANL flow sorted chromosome 5 library in | | | | | | | | |
| | M13mp18" | | | | | | | | |
| | /note="5qM" | | | | | | | | |
| STS | 33..139 | | | | | | | | |
| | /note="LANL STS 122" | | | | | | | | |
| | /complement(33..57) | | | | | | | | |
| | /note="for primer A" | | | | | | | | |
| primer_bind | /evidence=experimental | | | | | | | | |
| | 115..139 | | | | | | | | |
| primer_bind | /note="for primer B" | | | | | | | | |
| | /evidence=experimental | | | | | | | | |
| BASE COUNT | 39 a 36 c 46 g 57 t | | | | | | | | |
| ORIGIN | | | | | | | | | |
| Query Match | 12.8%; Score 23.6; DB 11; Length 178; | | | | | | | | |
| Best Local Similarity | 69.6%; Pred. No. 8.4e+03; | | | | | | | | |
| Matches 32: Conservative | 0; Mismatches 14; Indels 0; Gaps 0; | | | | | | | | |
| OY | 8 | TCTTCAGGTACAGACAGTGTGTGTCTGTGTGGCTAACCT | 53 | | | | | | |
| Db | 39 | TCTTACTGTACACACCTTGTATTATTCATCTTATGTCCCTAACCT | 84 | | | | | | |
| LOCUS | AU028588 | 149 bp | DNA | linear | STS 02-MAR-1999 | | | | |
| DEFINITION | Rattus norvegicus, OTSUKA clone, OT39.39/108c05, microsatellite | | | | | | | | |
| ACCESSION | AU028588 | | | | | | | | |
| VERSION | AU028588.1 | GI:4518511 | | | | | | | |
| KEYWORDS | STS. | | | | | | | | |
| SOURCE | Rattus norvegicus DNA, clone:OT39.39/108c05. | | | | | | | | |
| ORGANISM | Rattus norvegicus | | | | | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | | | | |
| | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; | | | | | | | | |
| | Rattus. | | | | | | | | |
| | 1 (sites) | | | | | | | | |
| REFERENCE | Matanabe,T.K., Hishigaki,H., Kanemoto,N., Miyakita-Mizoguchi,A., | | | | | | | | |
| AUTHORS | Oga,K., Okuno,S., Ono,T., Tsuji,A., Hayashi,H., Adachi,M., | | | | | | | | |
| | Yamasaki,Y., Iriye,Y., Takahashi,E., Takagi,T., Nakamura,Y. and | | | | | | | | |
| | Tanigami,A. | | | | | | | | |
| | The large-scale mapping of rat microsatellite markers | | | | | | | | |

| | |
|---|--|
| JOURNAL | Unpublished |
| REFERENCE | 2 (bases 1 to 149) |
| AUTHORS | Watanabe,T.K. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (24-JUL-1998) Takeshi K Watanabe, Otsuka Pharmaceutical Co., Ltd., Otsuka GEN Research Institute; 463-10, Kagasuno, Kanouchi-cho, Tokushima, Tokushima 771-0197, Japan (E-mail:watanabetsukaka.genome.ad.jp, Tel:81-886-65-2888, Fax:81-886-37-1035) |
| FEATURES | Location/Qualifiers |
| SOURCE | 1..149 |
| | /organism="Rattus norvegicus" |
| | /db_xref="taxon:10116" |
| | /clone="OT39_39/108c05" |
| | /note="OT39_39/108c05F-5'-GATGCTCCTGGGTACTTGA-3', OT39_39/108c05R-5'-AATTACATAGACAGCACACTAG-3'" |
| BASE COUNT | 26 a 18 c 44 g 60 t 1 others |
| ORIGIN | |
| Query Match | 12.7%; Score 23.4; DB 11; Length 149; |
| Best Local Similarity | 73.2%; Pred. No. 9.8e+03; |
| Matches | 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0; |
| OY | 1 ACTTCATTTCTTGACGTACAGACAGTGTGTGTTCTCGT 41 |
| Dn | 59 ACTGATTTGGAGSTACAGTAGTGTGTGTGTGTGTGTGT 99 |
| LOCUS | G19018 175 bp DNA linear STS 13-MAR-1996 |
| DEFINITION | cow STS BMS2172, sequence tagged site. |
| ACCSSION | G19018 |
| VERSION | G19018.1 GI:1223475 |
| KEYWORDS | STS; STS sequence; primer; sequence tagged site. |
| SOURCE | Bos taurus. |
| ORGANISM | Bos taurus |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos. 1 (bases 1 to 175) Stone,R.T. Cow Markers Unpublished (1996) | |
| CONTACT: | Roger Stone |
| U.S. Meat Animal Research Center | |
| P.O. Box 166, Clay Center, NE 68933 | |
| Tel: (402) 762-4166 | |
| Fax: (402) 762-4173 | |
| PRIMER A: | TTTTAGTGCCTCAATCCCTATG |
| PRIMER B: | TATTCACATTGTGCATTCATCC |
| STS size: | 80. |
| LOCATION/QUALIFIERS | 1..175 |
| /organism="Bos taurus" | |
| /db_xref="taxon:9913" | |
| STS | 31..53 |
| primer_bind | 31..53 |
| complement(88..110) | |
| BASE COUNT | 50 a 19 c 45 g 60 t 1 others |
| ORIGIN | |
| Query Match | 12.7%; Score 23.4; DB 11; Length 175; |
| Best Local Similarity | 55.6%; Pred. No. 9.7e+03; |
| Matches | 45; Conservative 0; Mismatches 36; Indels 0; Gaps 0; |
| OY | 24 GTGTTGTTGTTCTCTGTTGGCTAACTTTAAGAAGCGGTCTTCAGAAATGATAAG 83 |
| Dn | 67 GTGTTGTTGTTGTTGTTGTAATGAACAAGAGCATTAATCTGAATAATMAAAT 126 |

| | |
|-----------------------|---|
| RESULT | 13 |
| LOCUS | AF088775/c |
| DEFINITION | Homo sapiens clone TCRBV20S1.82 T-cell receptor beta chain (TCRBV20S1) mRNA, partial cds. |
| ACCESSION | AF088775 |
| VERSION | AF088775.1 GI:4038240 |
| KEYWORDS | |
| SOURCE | |
| ORGANISM | Homo sapiens. |
| REFERENCE | Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |
| AUTHORS | 1 (bases 1 to 112) |
| TITLE | Manifras,B.J. |
| JOURNAL | Rearranged human TCRBV20S1 genes |
| REFERENCE | Unpublished |
| AUTHORS | 2 (bases 1 to 112) |
| TITLE | Manifras,B.J. |
| JOURNAL | Direct Submission |
| FEATURES | Submitted (31-AUG-1998) Department of Internal Medicine, University Hospital Ulm, Robert-Koch-Str. 8, Ulm D-89081, Germany |
| SOURCE | location/Qualifiers |
| gene | 1..112 /organism="Homo sapiens" |
| CDS | /db_xref="taxon:9606" /clone="TCRBV20S1.82" /cell_type="T-lymphocyte" <1..>112 /gene="TCRBV20S1" <1..>112 /gene="TCRBV20S1" /codon_start=1 /product="T-cell receptor beta chain" /protein_id="AAC97324.1" /db_xref="GI:4038241" /translation="LSDSGFYLCAWSYORPLPHRGNGTRLTYEDLNKVP" |
| BASE COUNT | 23 a 34 c 29 g 26 t |
| ORIGIN | |
| Query Match | 12.4%; Score 22.8; DB 9; Length 112; |
| Best Local Similarity | 62.1%; Pred. No. 1.6e+04; |
| Matches | 36; Conservative 0; Mismatches 22; Indels 0; Gaps 0; |
| Oy | 118 GTAACCTGTGCGCGTTCCATTGCACAGAATCAGAGCCTCATTCCTTAGTAGTG 175 |
| Db | 109 GGAAACACTGTGTGAGGCTCTGTCGACAGATGAGGCTGTGCCATTCCCAAAGTGAG 52 |
| RESULT | 14 |
| LOCUS | AF088796/c |
| DEFINITION | Homo sapiens clone TCRBV20S1.103 T-cell receptor beta chain (TCRBV20S1) mRNA, partial cds. |
| ACCESSION | AF088796 |
| VERSION | AF088796.1 GI:4038281 |
| KEYWORDS | |
| SOURCE | |
| ORGANISM | Homo sapiens. |
| REFERENCE | Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |
| AUTHORS | 1 (bases 1 to 112) |
| TITLE | Manifras,B.J. |
| JOURNAL | Rearranged human TCRBV20S1 genes |
| REFERENCE | Unpublished |
| AUTHORS | 2 (bases 1 to 112) |
| TITLE | Manifras,B.J. |
| JOURNAL | Direct Submission |
| FEATURES | Submitted (31-AUG-1998) Department of Internal Medicine, University Hospital Ulm, Robert-Koch-Str. 8, Ulm D-89081, Germany |
| SOURCE | location/Qualifiers |
| gene | 1..112 /organism="Homo sapiens" |
| CDS | /db_xref="taxon:9606" /clone="TCRBV20S1.82" /cell_type="T-lymphocyte" <1..>112 /gene="TCRBV20S1" <1..>112 /gene="TCRBV20S1" /codon_start=1 /product="T-cell receptor beta chain" /protein_id="AAC97324.1" /db_xref="GI:4038241" /translation="LSDSGFYLCAWSYORPLPHRGNGTRLTYEDLNKVP" |
| BASE COUNT | 23 a 34 c 29 g 26 t |
| ORIGIN | |

| | | |
|-----------------------|---|---|
| | gene | /organism="Homo sapiens" |
| | CDS | <db_xref="taxon:9606"> /clone="TCRBV20S1.103" /cell_type="T-Lymphocyte" |
| | | <1..>112 /gene="TCRBV20S1" |
| | | <1..>112 /gene="TCRBV20S1" |
| | | /codon_start=1 /product="T-cell receptor beta chain" |
| | | /protein_id="AAC97344.1" |
| | | /db_xref="GI:4038126" |
| | | /translation="LSDSGFYLCAMSVQRPPLHFGNGTRLTVTEDLNKVP" |
| BASE COUNT | 23 a | 34 c |
| ORIGIN | | 29 g |
| | | 26 t |
| Query Match | 12.4% | Score 22.8; DB 9; Length 112; |
| Best Local Similarity | 62.1% | Pred. No. 1.6e+04; |
| Matches | 36; Conservative | 0; Mismatches 22; Indels 0; Gaps 0; |
| OY | 118 | GTAACCCGTGTCGGCGCTCATATGCACAGAAGTACAGCCTCATTCCTAGTAGTG 175 |
| Db | 109 | GAACACCTTTGTCAGGCTCTGTGTCACAGTAGAGCCTGTCCATTCCCAAAAGTGAG 52 |
| RESULT 15 | AF088715/c | 121 bp mRNA linear PRI 19-DEC-1998 |
| LOCUS | Homo sapiens clone TCRBV20S1.22 T-cell receptor beta chain | |
| DEFINITION | (TCRBV20S1) mRNA, partial cds. | |
| ACCESSION | AF088715 | |
| VERSION | AF088715.1 GI:4038125 | |
| KEYWORDS | . | |
| SOURCE | Homo sapiens. | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| | Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo. | |
| REFERENCE | 1 (bases 1 to 121) | |
| AUTHORS | Manifras,B.J. | |
| JOURNAL | Rearranged human TCRBV20S1 genes | |
| REFERENCE | Unpublished | |
| AUTHORS | 2 (bases 1 to 121) | |
| TITLE | Manifras,B.J. | |
| JOURNAL | Direct Submission | |
| | Submitted (31-AUG-1998) Department of Internal Medicine, University | |
| | Hospital Ulm, Robert-Koch-Str. 8, Ulm D-89081, Germany | |
| FEATURES | location/Qualifiers | |
| source | 1..121 | |
| | /organism="Homo sapiens" | |
| | /db_xref="taxon:9606" | |
| | /clone="TCRBV20S1.22" | |
| | /cell_type="T-Lymphocyte" | |
| | <1..>121 | |
| gene | /gene="TCRBV20S1" | |
| CDS | <1..>121 | |
| | /gene="TCRBV20S1" | |
| | /codon_start=1 | |
| | /product="T-cell receptor beta chain" | |
| | /protein_id="AAC97269.1" | |
| | /db_xref="GI:4038126" | |
| | /translation="LSDSGFYLCAMSVQSLNSPILHFNGTRLTVTEDLNKVP" | |
| BASE COUNT | 24 a | 35 c |
| ORIGIN | | 31 t |
| Query Match | 12.4% | Score 22.8; DB 9; Length 121; |
| Best Local Similarity | 62.1% | Pred. No. 1.6e+04; |
| Matches | 36; Conservative | 0; Mismatches 22; Indels 0; Gaps 0; |
| OY | 118 | GTAACCCGTGTCGGCGCTCATATGCACAGAAGTACAGCCTCATTCCTAGTAGTG 175 |
| Db | 118 | GAACACCTTTGTCAGGCTCTGTGTCACAGTAGAGCCTGTCCATTCCCAAAAGTGAG 61 |

Wed Feb 19 10:49:12 2003

us-09-939-293-1_copy_56_239.lim.rge

Page 7

Search completed: February 16, 2003, 07:15:19
Job time : 3215 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 05:03:52 ; Search time 296 Seconds

(without alignments)
1399.891 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 184
Sequence: 1 actcattcttcagtgacag.....ccttagtagtaagattga 184

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2583070

Minimum DB seq length: 0
Maximum DB seq length: 184

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : N_Geneseq_101002:*

- 1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 25.6 | 13.9 | 162 | 22 | ABAT72770 |
| 2 | 25.6 | 13.9 | 162 | 22 | ABA38410 |
| 3 | 25.6 | 13.9 | 162 | 22 | AAK21199 |
| 4 | 25.6 | 13.9 | 162 | 22 | AAK47356 |
| 5 | 25.6 | 13.9 | 162 | 22 | AAI25917 |
| 6 | 25.6 | 13.9 | 162 | 22 | AAI53192 |
| 7 | 25.6 | 13.9 | 162 | 22 | AB521511 |
| 8 | 25.6 | 13.6 | 158 | 21 | AAI30430 |
| 9 | 24.8 | 13.5 | 148 | 24 | ABL71429 |

| | | | | | | | |
|---|----|------|------|-----|----|-----------|--------------------|
| C | 10 | 24.4 | 13.3 | 163 | 20 | AAH85700 | Human single nucle |
| C | 11 | 23.8 | 12.9 | 114 | 24 | ABK79881 | Bacillus clausii g |
| C | 12 | 23.6 | 12.9 | 181 | 23 | AA482855 | Enterococcus faeca |
| C | 13 | 23.6 | 12.8 | 169 | 21 | AAI31534 | Human secreted pro |
| C | 14 | 23.4 | 12.7 | 152 | 24 | AB523848 | Human genome-deriv |
| C | 15 | 23.2 | 12.6 | 154 | 16 | AAI23779 | Human gene signatu |
| C | 16 | 23.2 | 12.6 | 174 | 18 | AAI80647 | Type II topoisomer |
| C | 17 | 22.8 | 12.4 | 157 | 21 | AAI08166 | Human secreted pro |
| C | 18 | 22.8 | 12.4 | 164 | 14 | AAI06518 | Human brain expres |
| C | 19 | 22.4 | 12.2 | 148 | 22 | AAK59274 | Human immune/haema |
| C | 20 | 22.2 | 12.1 | 157 | 22 | AAK20331 | Human brain expres |
| C | 21 | 22.2 | 12.1 | 157 | 24 | AB520737 | Human genome-deriv |
| C | 22 | 22.2 | 12.1 | 169 | 22 | AAH70328 | Human cervical can |
| C | 23 | 22.2 | 12.1 | 172 | 21 | AAI16632 | Human secreted pro |
| C | 24 | 22 | 12.0 | 125 | 22 | AAI90658 | Human digestive sy |
| C | 25 | 22 | 12.0 | 132 | 21 | AAI15687 | Human prostate can |
| C | 26 | 22 | 12.0 | 132 | 22 | AAI72789 | Human genome-deriv |
| C | 27 | 22 | 12.0 | 171 | 23 | AAI483029 | Enterococcus faeca |
| C | 28 | 21.8 | 11.8 | 114 | 22 | AAI23121 | Human brain expres |
| C | 29 | 21.8 | 11.8 | 131 | 16 | AAI24773 | Human gene signatu |
| C | 30 | 21.8 | 11.8 | 163 | 22 | ABAT6684 | Human foetal liver |
| C | 31 | 21.8 | 11.8 | 163 | 22 | ABAT4187 | Probe #19653 for g |
| C | 32 | 21.8 | 11.8 | 163 | 22 | AAK25316 | Human brain expres |
| C | 33 | 21.8 | 11.8 | 163 | 22 | AAK51323 | Human bone marrow |
| C | 34 | 21.8 | 11.8 | 163 | 22 | AAI28334 | Probe #18267 for g |
| C | 35 | 21.8 | 11.8 | 163 | 22 | AAI57381 | Probe #26067 used |
| C | 36 | 21.8 | 11.8 | 163 | 24 | AB524866 | Human genome-deriv |
| C | 37 | 21.8 | 11.8 | 177 | 16 | AAI24123 | Human gene signatu |
| C | 38 | 21.4 | 11.6 | 105 | 24 | ABL83354 | Human genome-deriv |
| C | 39 | 21.4 | 11.6 | 126 | 18 | AAI75584 | Staphylococcus aur |
| C | 40 | 21.4 | 11.6 | 170 | 22 | AAI35900 | Human genome-deriv |
| C | 41 | 21.2 | 11.5 | 172 | 24 | AB517770 | Human genome-deriv |
| C | 42 | 21.2 | 11.5 | 154 | 22 | AAI90341 | Peanut allergen Ar |
| C | 43 | 21.2 | 11.5 | 180 | 24 | ABAT7274 | Human genome-deriv |
| C | 44 | 21.2 | 11.5 | 184 | 16 | AAI23603 | Human gene signatu |
| C | 45 | 21 | 11.4 | 153 | 22 | ABAT4607 | Human foetal liver |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| ID | ABAT72770 standard; DNA; 162 BP. |
| AC | ABAT72770; |
| XX | |
| DT | 01-FEB-2002 (first entry) |
| XX | |
| DE | Human foetal liver single exon nucleic acid probe #21075. |
| XX | |
| KW | Human; foetal liver; gene expression; single exon nucleic acid probe: ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200157277-A2. |
| XX | |
| PD | 09-AUG-2001. |
| XX | |
| PF | 30-JAN-2001; 2001WO-US00669. |
| XX | |
| PR | 04-FEB-2000; 2000US-0180312. |
| PR | 26-MAY-2000; 2000US-0207456. |
| PR | 30-JUN-2000; 2000US-0608408. |
| PR | 03-AUG-2000; 2000US-0632366. |
| PR | 21-SEP-2000; 2000US-0234687. |
| PR | 27-SEP-2000; 2000US-0236359. |
| PR | 04-OCT-2000; 2000US-0024263. |
| XX | |
| PA | (MOL-) MOLECULAR DYNAMICS INC. |
| XX | |
| PI | Penn SG, Hanzel DK, Chen W, Rank DR; |
| XX | |

DR WPI: 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 4; SEQ ID NO 21075: 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 162 BP; 53 A; 37 C; 34 G; 38 T; 0 other;
XX
Query Match 13.9%; Score 25.6; DB 22; Length 162;
Best Local Similarity 55.7%; Pred. No. 70;
Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
XX
QY 7 TTCTTCAGGTACAGACAGTGTGTGTCTCTGTGCGTAACTTTAAGACGGGT 66
DB 65 TTCTTCACATTCAGACAGTGTCTCAATCTGAGTGAAGAACCACTATGACTCTGTCT 124
XX
QY 67 TTCTCAGAAATGATTAAGACCATGCGACA 94
DB 125 GGCTCACACTTGCTTAAGACCAAGCCATA 152
XX
RESULT 2
ABA38410
ID ABA38410 standard; DNA; 162 BP.
XX
AC ABA38410;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #16876 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 4; SEQ ID No 16876; 530pp; English.
XX

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 162 BP; 53 A; 37 C; 34 G; 38 T; 0 other;
XX
Query Match 13.9%; Score 25.6; DB 22; Length 162;
Best Local Similarity 55.7%; Pred. No. 70;
Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
XX
QY 7 TTCTTCAGGTACAGACAGTGTGTGTCTCTGTGCGTAACTTTAAGACGGGT 66
DB 65 TTCTTCACATTCAGACAGTGTCTCAATCTGAGTGAAGAACCACTATGACTCTGTCT 124
XX
QY 67 TTCTCAGAAATGATTAAGACCATGCGACA 94
DB 125 GGCTCACACTTGCTTAAGACCAAGCCATA 152
XX
RESULT 3
AAK21199
ID AAK21199 standard; DNA; 162 BP.
XX
AC AAK21199;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 21190.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 21190; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.

XX Sequence 162 BP; 53 A; 37 C; 34 G; 38 T; 0 other;

Query Match 13.9%; Score 25.6; DB 22; Length 162;

Best Local Similarity 55.7%; Pred. No. 70; Mismatches 39; Indels 0; Gaps 0;

Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 7 TTCTTCAGGTACAGACAGTGTGTGCTTCCTGTTGGCTTAACCTTTAAGACGGGTG 66

DB 65 TTCTTCACATTCAGACAGTGTGCTCAATCTGAGTGGAGAAGAACCACTATGAACCTTCTCTG 124

QY 67 TTCTCAGAAATTGATTAAGACCATGGCACA 94

DB 125 GGGTCACACTTGTCTAAGCAAGCCATA 152

RESULT 4

AAK47356

ID AAK47356 standard; DNA; 162 BP.

XX AAK47356;

AC AAK47356;

XX 06-NOV-2001 (first entry)

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 21913.

DE Human bone marrow expressed single exon probe SEQ ID NO: 21913.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukemia; lymphoma; myeloma; ss.

XX Homo sapiens.

OS Homo sapiens.

XX WO200157276-A2.

PN WO200157276-A2.

XX 09-AUG-2001.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488900/53.

XX WPI; 2001-488900/53.

DR WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 21913; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is one of

CC the probes of the invention.

XX Sequence 162 BP; 53 A; 37 C; 34 G; 38 T; 0 other;

XX Query Match 13.9%; Score 25.6; DB 22; Length 162;

XX Best Local Similarity 55.7%; Pred. No. 70;

XX Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 7 TTCTTCAGGTACAGACAGTGTGTGCTTCCTGTTGGCTTAACCTTTAAGACGGGTG 66

DB 65 TTCTTCACATTCAGACAGTGTGCTCAATCTGAGTGGAGAAGAACCACTATGAACCTTCTCTG 124

QY 67 TTCTCAGAAATTGATTAAGACCATGGCACA 94

DB 125 GGGTCACACTTGTCTAAGCAAGCCATA 152

DB 65 TTCTTCACATTCAGACAGTGTGCTCAATCTGAGTGGAGAAGAACCACTATGAACCTTCTCTG 124

QY 67 TTCTCAGAAATTGATTAAGACCATGGCACA 94

DB 125 GGGTCACACTTGTCTAAGCAAGCCATA 152

RESULT 5

AAI25917

ID AAI25917 standard; DNA; 162 BP.

XX AAI25917;

AC AAI25917;

XX 12-OCT-2001 (first entry)

DT 12-OCT-2001 (first entry)

DE Probe #15850 for gene expression analysis in human cervical cell sample.

XX Probe; human; microarray; gene expression; cervical epithelial cell.

KW cervical cancer; ss.

XX Homo sapiens.

OS Homo sapiens.

XX WO200157278-A2.

PN WO200157278-A2.

XX 09-AUG-2001.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00670.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488901/53.

XX WPI; 2001-488901/53.

DR WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

XX Claim 25; SEQ ID NO 15850; 487bp; English.

PS The present invention relates to human single exon nucleic acid probes

CC (SENP). The present sequence is one such probe. The SENPs are derived

CC from human HeLa cells. The SENPs can be used to produce a single exon

CC microarray, which can be used for measuring human gene expression in a

CC sample derived from human cervical epithelial cells. By measuring gene

CC expression, the probes are therefore useful in grading and/or staging

CC of diseases of the cervix, notably cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 162 BP; 53 A; 37 C; 34 G; 38 T; 0 other;

XX Query Match 13.9%; Score 25.6; DB 22; Length 162;

XX Best Local Similarity 55.7%; Pred. No. 70;

XX Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 7 TTCTTCAGGTACAGACAGTGTGTGCTTCCTGTTGGCTTAACCTTTAAGACGGGTG 66

DB 65 TTCTTCACATTCAGACAGTGTGCTCAATCTGAGTGGAGAAGAACCACTATGAACCTTCTCTG 124

QY 67 TTCTCAGAAATTGATTAAGACCATGGCACA 94

DB 125 GGGTCACACTTGTCTAAGCAAGCCATA 152

RESULT 6
AA153192
ID AA153192 standard; DNA; 162 BP.
XX
AC AA153192;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #21878 used to measure gene expression in human placenta sample.
XX
KW Probe: microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PT WPI; 2001-488897/53.
XX
DR Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
PT Claim 25; SEQ ID No 21878; 654bp; English.
XX
PS The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SO Sequence 162 BP; 53 A; 37 C; 34 G; 38 T; 0 other;
XX
Query Match 13.9%; Score 25.6; DB 22; Length 162;
Best Local Similarity 55.7%; Pred. No. 70;
Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
XX
QY 7 TTCTTCAGAGTACAGACAGCTTTGTCCTGTCCTGTCGTAACCTTAAGAGCGGCT 66
DB 65 TTCTTCAGATTCAGACAGCTGTCCTCAATCTGAGTGAAGACCACTATGAACTTCTGCTGT 124
QY 67 TTCTTCAGATTCAGATGAGACCATGGCACA 94
DB 125 GGGTCACACTTGCTTAAGACCAAGCATA 152
XX
RESULT 7
ABS21511
ID ABS21511 standard; DNA; 162 BP.
XX
AC ABS21511;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 21502.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberculous sclerotic; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PT WPI; 2002-114183/15.
XX
DR Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX
PS Claim 4; SEQ ID No 21502; 634bp; English.
XX
PT The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridize at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberculous sclerotic, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 13.5%; | Score 24.8; | DB 24; | Length 148; |
| Best Local Similarity | 57.9%; | Pred. No. 1.3e+02; | | |
| Matches 44; | Conservative 0; | Mismatches 32; | Indels 0; | Gaps 0; |

OY 92 ACACAACTGTGACGATTTGGCTTTGGAGTAACCCGTGTGGCGTTCCTATTGCACAGAAT 151
 || ||| ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 12 ACCAACAAGGCCAAGTAACTTTAGCAGAGCGTTATGTGCAGCTCCTGTAGAAATTCAA 71

QY 152 CAGAGCCTCATTCCT 167
||| ||| |||
Db 72 TTGAACCGTATTTCCT 87

RESULT 10
AAH85700/c
ID AAH85700 standard; DNA; 163 BP.

DT 27-FEB-2002 (first entry)

DE Human single nucleotide polymorphism containing DNA sequence #5557.

KM Biallelic marker; polymorphism; human; disease; diagnosis; treatment
 KM phenotypic trait; gene therapy; forensic; paternity; mapping; cancer
 KM transgenic; single nucleotide polymorphism; SNP; ss.

Homo sapiens.

| FH | Key | Location/Qualifiers |
|----|-----------|---------------------|
| FT | Variation | replace(21,G) |

```
FT /standard_name= "single nucleotide polymorphism"
```

PN W09953095-A2.

PD 21-OCT-1999.

PF 30-MAR-1999; 99WO-US06893.

PR 09-APR-1998; 98US-0057871.

PA (WHED) WHITEHEAD INST BIOMEDICAL RES

PI Lander ES, Wang D, Hudson T:

DR WPT; 1999-620443/53.

PT Polymorphic human genomic sequences and related allele-specific probes
PT and primers, useful for genetic analysis, e.g. diagnosis and monitoring
PT of disease -

PS Claim 1; Page 90; 330pp; English.

CC This invention describes novel human nucleic acid segments (I)
CC containing polymorphic sites. The polynucleotides of (I) are used for
CC e.g. correlating disease polymorphisms (or disease susceptibility) or
CC other phenotypic traits (e.g. baldness, obesity, fertility, strength,
CC response to drugs etc.); diagnosing and monitoring e.g. cancer,
CC inflammation, heart or central nervous system diseases; detecting
CC susceptibility to microbial infection; treating or preventing such
CC diseases; forensic analysis; gene therapy; paternity testing; mapping
CC genomic loci associated with phenotypic traits (and subsequent cloning
CC of the genes responsible); and the production of transgenic organisms.
CC Antibodies raised against (I) are useful as diagnostic and therapeutic
CC tools and in drug screening. AAH85144 - AAH87644 represent the human
CC DNA sequences containing biallelic polymorphic sites described in the
CC invention.

Sequence 163 BP; 45 A; 37 C; 30 G; 51 T; 0 other;

Query Match 13.3%; Score 24.4; DB 20; Length 163;

| | | |
|-----------------------|------------------|------------------------------------|
| Best Local Similarity | 53.1%; | Pred. No. 1.8e+02; |
| Matches | 52; Conservative | 0; Mismatches 46; Indels 0; Gaps 0 |

Oy 38 CTGTTGTGGCTTAACCTTAAAGAACGGTCITTTCTCAGAATTGATTAAGACCATTGGCAAAAA 97
||| ||| ||| | ||| | ||| ||| |||
Db 98 CTGCTTTGGGTAAACTGAGCAGAAAGTGATACACAGAAAGGAAAATGTGCACATCTATGCTA 39

qy 98 CTGTGACGATTTGGCCTTTGGAGTAAACCCCTGTGTGCGGTT 1
 |||| | | | | | | | | | |
 Db 38 GTGTGAATTTGGTAACTTGCCTGACCTCTGCAAGGCTGTT 1

RESULT 11
ABK79881/c
ID ABK79881 standard; DNA; 114 BP.

DT 13-AUG-2002 (first entry)

DE *Bacillus clausii* genomic sequence tag (GST) #2724.

KW Differential gene expression; genomic sequenced tag; GST;

KW physiological provocation; ds.

OS *Bacillus clausii*

PN WO200229113-A2

PD 11-APR-2002.

PF 05-OCT-2001; 2001WO-US31437.

PR 06-OCT-2000; 2000US-0680598.

[illegible]

PA (NOVO) NOVOZYMES AS.

PI Berka R, Clausen IG;

DR WPI; 2002-416684/44.

| PT | Monitoring differen |
|-------|---------------------|
| 25.25 | |

PT Bacillus cells, by using

XX

PS Claim 11; SEQ ID NO 7172; 200pp; English.

CC The invention describes a method for monitoring differential expression of
CC genes in a first *Bacillus* cell relative to expression of the genes in
CC other *Bacillus* cells, comprising hybridising labelled nucleic acid probes
CC isolated from *Bacillus* cells to a substrate containing array of *Bacillus*
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first *Bacillus* cell relative to expression of the same genes
CC in one or more second *Bacillus* cells. The method is useful for monitoring
CC global expression of several genes from a *Bacillus* cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which *Bacillus* cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive
CC follow-up characterisation is unnecessary, when one spot on an array
CC equals one gene or one open reading frame, since sequence information is
CC available. This sequence represents a genomic sequence tag (GST) used in
CC the method of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at

CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 114 BP; 34 A; 25 C; 26 G; 29 T; 0 other;
SQ

Query Match 12.9%; Score 23.8; DB 24; Length 114;
Best Local Similarity 57.3%; Pred. No. 2.5e+02;
Matches 43; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

YY 85 CCATGCGACAAAACGTGCGATGCTTGAGTAACCTGTGCGGCTTCCTATGCA 144
DB 96 CCATGCGCCCAATAAATGCTATGCTGTGCGCGACGCTGAATGAATGACCATGGCG 37
YY 145 CAGAAATCAGAGCCT 159
DB 36 GTGAATGACAGCCT 22

RESULT 12
ID AAS48255 standard; DNA; 181 BP.
XX
AC AAS48255;
XX
DT 13-FEB-2002 (first entry)
XX
DE Enterococcus faecalis cellular proliferation inhibitory sequence #825.
XX
KM Antisense: ss; prokaryotic cellular proliferation;
XX antibiotic; antibacterial; drug design.
XX
OS Enterococcus faecalis.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
XX 23-MAY-2000; 2000US-206848P.
XX 26-MAY-2000; 2000US-207727P.
XX 23-OCT-2000; 2000US-242578P.
XX 27-NOV-2000; 2000US-253625P.
XX 22-DEC-2000; 2000US-257931P.
XX 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DK WPI; 2001-611495/70.
XX
XX
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 1; Seq ID No 832; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 181 BP; 51 A; 28 C; 36 G; 66 T; 0 other;
XX

Query Match 12.9%; Score 23.8; DB 23; Length 181;
Best Local Similarity 55.4%; Pred. No. 3e+02; 37; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

YY 26 GTTGTGTCCTTCCTGTCGCTTAAGAGCGGTTCACGAATTGATAGAC 85
DB 50 GATATTTCCTTGAAATTCGTAATTAATCAAGGTAATGCTTGTAGAAATGATTTCAT 109
YY 86 CATGGCACAAAACGTGACGATT 108
DB 110 CACGAATTAAATCTGTGAAGGTT 132

RESULT 13
ID AAC31534/C
XX AAC31534 standard; cDNA; 169 BP.
XX
AC AAC31534;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 35609.
XX
KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI; 2000-500381/45.
XX
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX
PS Claim 1; SEQ ID 35609; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 169 BP; 41 A; 56 C; 32 G; 38 T; 2 other;
XX

Query Match 12.8%; Score 23.6; DB 21; Length 169;

Best Local Similarity: 53.4%; Pred. No. 3.4e+02;
Matches 47; Conservative 1; Mismatches 40; Indels 0; Gaps 0

OY 52 TTTAAGAAAGCGGTGTTTCTCAGAATTGATTAAGACCATGGCACAAAACGTGACGATTGCC 111
| | | | | | | | | : | | | | | | | | | | | | | | | |
Db 151 TTTAAAAAAGAGGCGSCSTCCAGAACAAATGTTGTTCAGTTCGTGAGAGTTAAAGT 92

OY 112 TTTTGAGTAACCCGTGTCGGGTCCOTA 139
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 91 GGTTAATAAATCTCTGCCTCTGCTCCGA 64

RESULT 14
ABS23848
ID ABS23848 standard; DNA; 152 BP.
XX AC ABS23848;
XX DT
DE 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon probe ORF from lung SEQ ID No 23839.

KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberosclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosterosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-180312P.
XX PR 26-MAY-2000; 2000US-207456P.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0633366.
XX PR 21-SEP-2000; 2000US-234687P.
XX PR 27-SEP-2000; 2000US-236359P.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPL; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
measure gene expression in human lung samples -
XX
PS Claim 4; SEQ ID No 23839; 634bp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
nucleic acid probes for measuring gene expression in a sample derived
from human lung comprising single exon nucleic acid probes having one
of 12614 nucleic acid sequences mentioned in the specification, or their
complements or the 12387 open reading frames derived from the 12614
probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
nucleic acid expressed in the human lung; measuring gene expression in a
sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
mRNA, and (b) measuring the label detectably bound to each probe of
the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
of the eukaryote; and (b) detecting specific hybridisation of detectably

| | |
|-----------|--|
| CC | labelednucleic acids from eukaryote lung mRNA,to a single exon probe, |
| CC | having a fragment identical to the predicted exon, the probe is included |
| CC | in the above mentioned microarray; assigning exons to a single gene, |
| CC | comprising (a) identifying exons from genomic sequence by the method |
| CC | above and (b) measuring the expression of each of the exons in several |
| CC | tissues and/or cell types using hybridisation to a single exon |
| CC | microarrays having a probe with the exon, where a common pattern of |
| CC | expression of the exons in the tissues and/or cell types indicates that |
| CC | the exons should be assigned to a single gene; a peptide comprising one |
| CC | of 1201 sequences, mentioned in the specification, or encoded by the |
| CC | probes/open reading frames (ORF). The probes are used for gene |
| CC | expression analysis, and for identifying exons in a gene, particularly |
| CC | using human lung derived mRNA and for the study of lung diseases |
| CC | such as asthma, lung cancer, chronic obstructive pulmonary disease |
| CC | (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary |
| CC | fibrosis, neurofibromatosis, tuberous sclerosis, gaucher's disease, |
| CC | Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary |
| CC | haemostederosis, pulmonary histiocytosis, lymphangioleiomyomatosis, |
| CC | pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic |
| CC | pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension |
| CC | and hyaline membrane disease. The present sequence is a single exon |
| CC | probe open reading frame of the invention. |
| CC | Note: The sequence data for this patent did not form part |
| CC | of the printed specification, but was obtained in electronic |
| CC | format directly from WIPO at |
| CC | ftp.wipo.int/pub/published_pcl_sequences. |
| XX | |
| SQ | Sequence 152 BP; 33 A; 40 C; 49 G; 30 T; 0 other; |
| | |
| | Query Match 12.7%; Score 23.4; DB 24; Length 152; |
| | Best Local Similarity 55.6%; Pred. No.3.8e+02; |
| Matches | 45; Conservative 0; Mismatches 36; Indels 0; Gaps 0; |
| OY | 75 ATTGTAAGACCATGCACAAATACTGTGACGATTGGCTTTGGATGAACCTGTGTGCGGT 134 Db 21 ATTACTTTGATTATCAAGCCCCGAGCTGGGCCGAGTCTCCTTGAGAGCTGACATTGAGCGGA 80 |
| OY | 135 TCCTATTGCACAGAANAATGAGA 155 Db 81 TACCTGTGAACACGAGAGCAGA 101 |
| RESULT 15 | |
| AAT23779 | |
| ID | AAT23779 standard; CDNA to mRNA; 154 BP. |
| XX | |
| AC | AAT23779; |
| XX | |
| DT | 13-SEP-1996 (first entry) |
| XX | |
| DE | Human gene signature H0MG505674. |
| XX | |
| KW | Gene signature; messenger RNA; mRNA; relative abundance; frequency; human; cloning; mapping; non-biased library; diagnosis; detection; cell typing; abnormal cell function; ss. |
| KW | |
| OS | Homo sapiens. |
| XX | |
| PN | WO9514772-A1. |
| XX | |
| PD | 01-JUN-1995. |
| XX | |
| PF | 11-NOV-1994; 94WO-JP01916. |
| XX | |
| PR | 12-NOV-1993; 93JP-0355504. |
| XX | |
| PA | (MATS/) MATSUBARA K. (OKUB/) OKUBO K. |
| XX | |
| PI | Matsubara K, Okubo K; |
| XX | |
| DR | WPI: 1995-206931/27. |
| XX | |

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues

PS Claim 1: Page 1453; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in AAT19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.

XX Sequence 154 BP; 50 A; 23 C; 22 G; 51 T; 8 other;

Query Match 12.6%; Score 23.2; DB 16; Length 154;

Best Local Similarity 61.7%; Pred No. 4.5e+02; Mismatches 23; Indels 0; Gaps 0;

OY 55 AAGAAAGCGGTGTTCTCAGAAATTGATAGACCATGCGACAAACTGTGACGATTGCTTT 114
 DB 67 AAGAGGAGCTGTTAATACATGATGATGATACCATGCGCAAAATCTTCTGAAATTGTCTT 126

Search completed: February 16, 2003, 06:21:33
 Job time : 298 secs


```

: LENGTH: 150 base pairs
: type: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: not relevant
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: US-08-189-256A-1

Query Match      11.1%, Score 20.4; DB 2; Length 150;
Best Local Similarity 53.8%; Pred. No. 3e+02;
Matches 42; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

OY 94 AAAATGTGAGCATGGCTTTGCAATAACCCGTGTGGCGTCCATATGACAGAARTCA 153
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 123 ACAACTGTATCAACAGCGCTTGCTATTCGCCCGAGATGCGTCCAGAAATATAGCATCC 64

OY 154 GAGCCTCATTCCTTAGT 171
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 63 CTGCCCCCTCAGCTCAT 46

RESULT 5
: US-09-193-853-1/c
: Sequence 1, Application US/09193853
: Patent No. 6388168
: GENERAL INFORMATION:
: APPLICANT: Maliga, Pal
: APPLICANT: Svab, Zora
: APPLICANT: Staub, Jeffrey
: APPLICANT: Zoubenko, Oleg V.
: APPLICANT: Allison, Lori A.
: APPLICANT: Kanevski, Ivan
: TITLE OF INVENTION: DNA Constructs and Methods for Stably
: TITLE OF INVENTION: Transforming Plasmids of Multicellular Plants and
: TITLE OF INVENTION: Expressing Recombinant Proteins Therein
: NUMBER OF SEQUENCES: 47
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Dann, Dorfman, Herrell and Skillman
: STREET: 1601 Market Street Suite 720
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103-2307
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/193,853
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/189,256
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/518,763
: FILING DATE: 01-MAY-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Reed, Janet E.
: REGISTRATION NUMBER: 36,252
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 563-4100
: TELEFAX: (215) 563-4044
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 150 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: not relevant
: MOLECULE TYPE: DNA (genomic)

```

```

;      HYPOTHETICAL: NO
;      ANTI-SENSE: NO
US-09-193-853-1

Query Match          11.1%; Score 20.4; DB 4; Length 150;
Best Local Similarity 53.8%; Pred. No. 3e+02;
Matches 42; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY   94  AAATGTCAGCATTTGGCTTGTGGACGTAAACCCTGTGTCCGCGTTCATTGCACAGAANTCA 153
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   123  ACAACTGTATCCACAACGCGCTTCGTAATGCCCGGAGTTCGCTCCAGAAATATAGCANTCC 64

QY   154  GAGCCTCATTCCTTAGT 171
      ||| | | | | | | | |
DB    63  CTGCCCCCTCACGTCAT 46

RESULT 6
US-08-189-256A-18/c
; Sequence 18, Application US/08189256A
; Patent No. 5877402
; GENERAL INFORMATION:
; APPLICANT: Maliga, Pal
; APPLICANT: Svab, Zora
; APPLICANT: Staud, Jeffrey
; APPLICANT: Zoubenko, Oleg V.
; APPLICANT: Allison, Lori A.
; APPLICANT: Carrier, Helaine
; APPLICANT: Kanevski, Ivan
; TIME OF INVENTION: DNA Constructs and Methods for Stably
; TITLE OF INVENTION: Transforming Plastids of Multicellular Plants and
; TITLE OF INVENTION: Expressing Recombinant Proteins Therein
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/189,256A
; FILING DATE: 31-JAN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/111,398
; FILING DATE: 25-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/518,763
; FILING DATE: 01-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-189-256A-18

Query Match          11.1%; Score 20.4; DB 2; Length 161;

```



```

RESULT 9
US-09-193-853-4/c
? Sequence 4, Application US/09193853
? Patent No. 6388168
? GENERAL INFORMATION:
? APPLICANT: Maliga, Pal
? APPLICANT: Svab, Zora
? APPLICANT: Staub, Jeffrey
? APPLICANT: Zoubenko, Oleg V.
? APPLICANT: Allison, Lori A.
? APPLICANT: Carreir, Helaine
? APPLICANT: Kanevski, Ivan
? TITLE OF INVENTION: DNA Constructs and Methods for Stably
? TITLE OF INVENTION: Transforming Plasmids of Multicellular Plants and
? TITLE OF INVENTION: Expressing Recombinant Proteins Therein
? NUMBER OF SEQUENCES: 47
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Dann, Dorfman, Herrell and Skillman
? STREET: 1601 Market Street Suite 720
? CITY: Philadelphia
? STATE: PA
? COUNTRY: USA
? ZIP: 19103-2307
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentln Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/193,853
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/189,256
? FILING DATE:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/518,763
? FILING DATE: 01-MAY-1990
? ATTORNEY/AGENT INFORMATION:
? NAME: Reed, Janet E.
? REGISTRATION NUMBER: 36,252
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (215) 563-4100
? TELEFAX: (215) 563-4044
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 165 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? HYPOTHEICAL: NO
? ANTI-SENSE: NO
US-09-193-853-4

Query Match      11.1%; Score 20.4; DB 4; Length 165;
Best Local Similarity 53.8%; Pred. No. 3.1e+02;
Matches 42; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY   94 AAACGTGAGCATTTGGATGAACCCTGTGCCTTCCATTATGACAGAATAACA 153
     | ||||| | | | | | | | | | | | | | | | | | | | | | |
Db    123 ACAACTGTAATCAAGCGCTTGTAATGCCCGCAGATTCTCCTCCAATAAATAGCATTC 64

QY   154 GAGCCTATTCCCTTAGT 171
     ||| | | | | | | | | | | | | | | | | | | | | | | | |
Db    63 CTGCCCTCACATCAAT 46

RESULT 10
US-08-189-256A-2/c
Sequence 2, Application US/08189256A

```

```

Patent No.5877402
GENERAL INFORMATION:
APPLICANT: Maliga, Pal
APPLICANT: Svab, Zora
APPLICANT: Staub, Jeffrey
APPLICANT: Zoubenko, Oleg V.
APPLICANT: Allison, Lori A.
APPLICANT: Carter, Helaine
APPLICANT: Kanevski, Ivan
TITLE OF INVENTION: DNA Constructs and Methods for Stably
TITLE OF INVENTION: Transforming Plasmids of Multicellular Plants and
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman
STREET: 1601 Market Street Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/189,256A
FILING DATE: 31-JAN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/111,398
FILING DATE: 25-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/518,763
FILING DATE: 01-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 168 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-189-256A-2

Query Match          11.1%; Score 20.4; DB 2; Length 168;
Best Local Similarity 53.8%; Pred.No.3.2e+02;
Matches 42; Conservative 0; Mismatches 36; Indels 0; Gaps 0.

QY      94  AAAACTGTGAGCATGGCTTTGGAGTAAACCGTGTCGGTTCCTATTCACAGAATAACA 153
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       123 ACAACTGTATCCAACCGCTTCTATTCCGCCGGAATTCCTCCCAATAATATAGCATCC 64
QY      154 GAGCCTCATTCCTCTAGT 171
        ||| | | | | | | |
Db       63 CTGCCCCCTCACGTCAT 46

RESULT 11
US-09-193-853-2/c
Sequence 2, Application US/09193853
Patent No. 6388168
GENERAL INFORMATION:
APPLICANT: Maliga, Pal
APPLICANT: Svab, Zora
APPLICANT: Staub, Jeffrey

```

```

;
; APPLICANT: Zoubenko, Oleg V.
; APPLICANT: Allison, Lori A.
; APPLICANT: Carrier, Helaine
; APPLICANT: Kanevski, Ivan
; TITLE OF INVENTION: DNA Constructs and Methods for Stably
; TITLE OF INVENTION: Transforming Plasmids of Multicellular Plants and
; TITLE OF INVENTION: Expressing Recombinant Proteins Therein
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/193,853
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/189,256
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/518,763
; FILING DATE: 01-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-09-193-853-2

Query Match 11.1%; Score 20.4; DB 4; Length 168;
Best Local Similarity 53.8%; Pred. No. 3.2e+02;
Matches 42; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 94 AAAAGCTGACGATGTGGAGTACCTGTGCGGTTCTATTGACAGAAATCA 153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 ACAAGCTGTATCCAAAGCGCTTCGTATTCGCCGAGTTCGCCAGAAATATAGCATCC 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 154 GAGCCTCATTCCTTAGT 171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 CTCGCCCTCAGCTCAAT 46

RESULT 12
US-09-424-620B-21
; Sequence 21, Application US/09424620B
; Patent No. 6391385
; GENERAL INFORMATION:
; APPLICANT: HANIL SYNTHETIC FIBER CO., LTD.
; JANG, Ki-Ryong
; MOON, Jae-Woong
; BAE, Cheon-Soon
; YANG, Doo-Suk
; LEE, Jee-Won
; SEONG, Baik-Lin
; TITLE OF INVENTION: Process for preparing recombinant proteins using highly
```

```

;
; efficient expression vector from Saccharomyces cerevisiae
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BACHMAN & LAPOINTE, P.C.
; STREET: Suite 1201, 900 Chapel Street
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: U.S.A.
; ZIP: 06510-2802
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM
; OPERATING SYSTEM: WINDOWS 95/98
; SOFTWARE: MS WORD
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/424,620B
; FILING DATE: 24-No. 6391585-1999
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: oligonucleotide
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-424-620B-21

Query Match 11.0%; Score 20.2; DB 4; Length 165;
Best Local Similarity 48.7%; Pred. No. 3.7e+02;
Matches 55; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 21 ACAAGTGTGTGTTCTCTGTGTGCTAATTAAAGCGGTGTTCTCAGAAATGAT 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 26 ACAATCTTACATCTCTCTGTGTGTTGTTGCTTCAAGTTGTCTACGTAAGTCA 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 81 AAGACCATGGCAGAAACCTGACGATGGCTTTGGAGTAACTGTGTCGG 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 86 CGTCCGGGACACAGCAAAAGCTGTGTGATGTCGTGCATATGAGACTG 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-08-569-147-68/C
; Sequence 68, Application US/08569147
; Patent No. 6180377
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMANISED ANTIBODIES
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/569,147
; FILING DATE: 25-March-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Iturjillo, Doreen Yalko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CARP-0047
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 68:
```

Search completed: February 16, 2003, 07:53:33
Job time : 69 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 07:15:22 ; Search time 99 Seconds

(Without alignments)
946.622 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 184
Sequence: 1 actcattctcttagtgtagcag.....ccttagtgtagtggaagattgga 184

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 325486

Minimum DB seq length: 0

Maximum DB seq length: 184

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Published Applications-NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 25.6 | 13.9 | 162 | 10 | US-09-864-761-23730 |
| 2 | 24.8 | 13.5 | 148 | 10 | US-09-294-0093B-803 |
| 3 | 23.8 | 12.9 | 114 | 10 | US-09-974-300-7172 |
| 4 | 23.4 | 12.9 | 181 | 10 | US-09-815-242-832 |
| 5 | 23.4 | 12.7 | 152 | 10 | US-09-864-761-33383 |
| 6 | 22.6 | 12.3 | 119 | 9 | US-09-994-228-82 |
| 7 | 22.2 | 12.1 | 157 | 10 | US-09-864-761-28343 |
| 8 | 22.2 | 12.0 | 132 | 10 | US-09-925-300-122 |
| 9 | 22.2 | 12.0 | 140 | 10 | US-09-783-590-1164 |
| 10 | 22.2 | 12.0 | 160 | 9 | US-09-728-444-1171 |
| 11 | 22.2 | 12.0 | 171 | 10 | US-09-815-242-786 |
| 12 | 21.8 | 11.8 | 114 | 10 | US-09-864-761-29692 |
| 13 | 21.8 | 11.8 | 163 | 10 | US-09-864-761-26507 |
| 14 | 21.8 | 11.8 | 163 | 10 | US-09-878-574-14495 |
| 15 | 21.8 | 11.8 | 167 | 9 | US-10-137-765-52 |
| 16 | 21.6 | 11.7 | 169 | 10 | US-09-969-373-1445 |
| 17 | 21.4 | 11.6 | 105 | 10 | US-09-867-701-6332 |
| 18 | 21.4 | 11.6 | 170 | 10 | US-09-764-869-1400 |
| 19 | 21.2 | 11.5 | 102 | 10 | US-09-864-761-32102 |

| | | | | | | |
|----|------|------|-----|----|---------------------|---------------------|
| 20 | 21.2 | 11.5 | 158 | 10 | US-09-878-574-14426 | Sequence 14426, A |
| 21 | 21.1 | 11.4 | 117 | 9 | US-09-728-444-324 | Sequence 324, App |
| 22 | 21.1 | 11.4 | 175 | 10 | US-09-864-761-29664 | Sequence 29664, A |
| 23 | 21.1 | 11.4 | 153 | 10 | US-09-864-761-18419 | Sequence 18419, A |
| 24 | 21.1 | 11.4 | 175 | 10 | US-09-815-242-772 | Sequence 772, App |
| 25 | 20.8 | 11.3 | 89 | 10 | US-09-864-761-24861 | Sequence 24861, A |
| 26 | 20.8 | 11.3 | 129 | 10 | US-09-864-761-25547 | Sequence 25547, A |
| 27 | 20.8 | 11.3 | 160 | 9 | US-10-040-739-498 | Sequence 498, App |
| 28 | 20.6 | 11.2 | 130 | 10 | US-09-728-446-440 | Sequence 440, App |
| 29 | 20.6 | 11.2 | 136 | 10 | US-09-783-590-6995 | Sequence 6995, App |
| 30 | 20.6 | 11.2 | 165 | 10 | US-09-969-373-457 | Sequence 457, App |
| 31 | 20.4 | 11.1 | 80 | 8 | US-08-961-888-21 | Sequence 21, Appl |
| 32 | 20.4 | 11.1 | 81 | 8 | US-08-961-888-21 | Sequence 21, Appl |
| 33 | 20.4 | 11.1 | 103 | 10 | US-09-864-761-19951 | Sequence 19951, A |
| 34 | 20.4 | 11.1 | 122 | 10 | US-09-783-590-2100 | Sequence 2100, App |
| 35 | 20.4 | 11.1 | 130 | 10 | US-09-783-590-7380 | Sequence 7380, App |
| 36 | 20.4 | 11.1 | 171 | 10 | US-09-864-761-28751 | Sequence 28751, A |
| 37 | 20.2 | 11.0 | 129 | 10 | US-09-867-701-3853 | Sequence 3853, App |
| 38 | 20.2 | 11.0 | 154 | 10 | US-09-563-817-481 | Sequence 481, App |
| 39 | 20.2 | 11.0 | 169 | 9 | US-09-754-853A-286 | Sequence 286, App |
| 40 | 20 | 10.9 | 135 | 10 | US-09-783-590-4094 | Sequence 4094, App |
| 41 | 20 | 10.9 | 184 | 10 | US-09-878-574-12896 | Sequence 12896, App |
| 42 | 19.8 | 10.8 | 118 | 10 | US-09-728-446-1453 | Sequence 1453, App |
| 43 | 19.8 | 10.8 | 125 | 10 | US-09-864-761-22217 | Sequence 22217, A |
| 44 | 19.8 | 10.8 | 141 | 10 | US-09-864-761-17657 | Sequence 17657, A |
| 45 | 19.8 | 10.8 | 178 | 10 | US-09-864-761-20835 | Sequence 20835, A |

ALIGNMENTS

RESULT 1
US-09-864-761-23730
; Sequence 23730, Application US/09864/61
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-x-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180, 312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207, 456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632, 366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263, 6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236, 359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

92 ACAAACTGTACGATTGGCTTTGAGTAACCCGTGTGCGTTCCATTGCACAGAAT 151

;
; PRIOR APPLICATION NUMBER: 60/253,625
;
; PRIOR FILING DATE: 2000-11-27
;

```
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 832
; LENGTH: 181
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-815-242-832

Query Match      12.9%; Score 23.8; DB 10; Length 181;
Best Local Similarity 55.4%; Pred. No. 1.2e+02;
Matches 46; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 26 GTTGTGCTTCTCTGTTGGCTAACTTTAAGAGCGGCTTTCTCAGATTGATAGAC 85
DB 50 GATATTGTTGTAATTTGTAATTAATTCAGAGTAATCGTTTGTGAAGATTTCAT 109
QY 86 CATGCACAAACTGTGACGATT 108
DB 110 CACGAATTAACTCTGTGAAGCTT 132

RESULT 5
US-09-864-761-32383
; Sequence 32383, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
```

```
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32383
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC022211.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7
; OTHER INFORMATION: NT HIT: 911432723, EVALU8 8.70e-02
; OTHER INFORMATION: SWISSPROT HIT: Q9SCW5, EVALU8 1.80e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE253522.1, EVALU8 2.00e-80
US-09-864-761-32383

Query Match      12.7%; Score 23.4; DB 10; Length 152;
Best Local Similarity 55.6%; Pred. No. 1.5e+02;
Matches 45; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 75 ATTGATTAAGACCATGSCACAAACTGTGACGANTGTGGCTTTGGAGTAAACCTGTGCGGT 134
DB 21 ATTACTTTGATTACTGCGCCCGAGCTGGCGGAGTCTCCCTGGAGCTGCACATTGAGCGGA 80
QY 135 TCCATTGCGACAGAAATCAGA 155
DB 81 TACCTCTGAACACCGACGACAGA 101

RESULT 6
US-09-994-228-82
; Sequence 82, Application US/09994228
; Publication No. US20030009016A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and M
; FILE REFERENCE: 21402-015 CIP
; CURRENT APPLICATION NUMBER: US/09/994,228
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 09/865,201
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/207,142
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-994-228-82

Query Match      12.3%; Score 22.6; DB 9; Length 119;
Best Local Similarity 58.0%; Pred. No. 2.5e+02;
Matches 40; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 87 ATGGCACAACACTGTGACGATTGGCTTTGGAGTAACCCCTGTGTGGCGTTCTATTGCACA 146
DB 49 ATGTGCACACACAGTACGATTGCTTTGGAGTAATCCAGATCTGAGAAAACAAACT 108
QY 147 GAATCAGA 155
DB 109 GAAGACAGA 117

RESULT 7
US-09-864-761-28343/c
; Sequence 28343, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
```

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aegmic-a-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28343
; LENGTH: 157
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006324.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.59
; OTHER INFORMATION: SWISSPROT HIT: Q90069, EVALUATE 4.60e-01
; OTHER INFORMATION: EST_HUMAN HIT: AI032107.1, EVALUATE 2.10e+00
; US-09-864-761-28343

Query Match          12.1%; Score 22.2; DB 10; Length 157;
Best Local Similarity 58.2%; Pred. No.3.7e+02;
Matches 39; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 28 TTGTGTCTCTGTTGGCTTAAGAAAGCGGTGTTTCTCAGATTGATTAAGACCA 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 124 TTGAGGGTCCCTGTTCTGGAGAACATATAGAACTGCACACACATCTGGGAGTCTT 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 88 TGGCACA 94
   ||| ||| |||
DB 64 AGAAACA 58

RESULT 8
; US-09-925-300-122/c
; Sequence 122, Application US/09925300
; Patent No. US20020151681A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 122
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-300-122

Query Match          12.0%; Score 22; DB 10; Length 132;
Best Local Similarity 55.1%; Pred. No. 4.1e+02;
Matches 43; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 41 TTGGCTTAACCTTTAAGAGCGGTGTTCTCAGATTGATTAAGACCAAACTG 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 129 TAGTGACAAAGTTGAAAGGCTGTGCTTTCCGTGATGTGACAGGATTAAGCG 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 101 TGACGATTGGCTTGAG 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 69 TCTCTGCTCCCTGTGGG 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
; US-09-783-590-1164
; Sequence 1164, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1164
; LENGTH: 140
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (108)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (121)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: (123)
; LOCATION: (123)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-1164

Query Match          12.0%; Score 22; DB 10; Length 140;
Best Local Similarity 54.4%; Pred. No. 4.2e+02;
```



```

PRIORITY APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 786
LENGTH: 171
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-815-242-786

Query Match          12.0%; Score 22; DB 10; Length 171;
Best Local Similarity 55.1%; Pred. No. 4.5e+02;
Matches 43; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY      26 GTTGTGCTTCCTCGTTGGCTAACTTAAGAACGGGTGTTCCACAATTGATAAAGC 85
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       93 GATATTGTTGGATTGATTAATCAAGTAATCGTTTGTAAGAATTGCAT 152
        || | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      86 CATGCCACAAACTGTGA 103
        || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       153 CACGATTAATTCTGTGA 170
        || | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-864-761-29692/c
Sequence 29692, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29692
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL162311.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.58
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.45
; OTHER INFORMATION: SWISSPROT HIT: P38343, EVALUATE 1.10e+00
; OTHER INFORMATION: NT HIT: AL163281.2, EVALUATE 4.00e-03
; OTHER INFORMATION: EST_HUMAN HIT: AA056128.1, EVALUATE 6.00e-03
US-09-864-761-29692
```

```

Query Match          11.8%; Score 21.8; DB 10; Length 114;
Best Local Similarity 61.4%; Pred. No. 4.5e+02;
Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
```

```

Qy      29 TGTGTGTCCTGTTGCTTACCTTTAGAGCGGTGTTCTCAGATTGATAGAC 85
Db      64 TGTGTTTCTCTCTCGGCTACCTTAAGATTCTTTTATCATGATTGATAC 8
```

```

RESULT 13
US-09-864-761-26507
; Sequence 26507, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26507
; LENGTH: 163
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006427.13
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EST_HUMAN HIT: AW903222.1, EVALUATE 6.00e-07
; OTHER INFORMATION: NT HIT: g11497621, EVALUATE 3.70e-01
US-09-864-761-26507
```

```

Query Match          11.8%; Score 21.8; DB 10; Length 163;
Best Local Similarity 56.2%; Pred. No. 5.1e+02;
Matches 41; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```

Qy      36 TCCCTGTTGCTTACCTTTAGAGCGGTGTTCTCAGATTGATAGACCATGACACA 95
Db      84 TCCAGCTCTGCGCTAACATCATGATATGATTCATGAGAGATCGATGACAGACACACA 143

Qy      96 AACCTGTGACGATT 108
Db      144 AGCTTAGACACTT 156
```

```

RESULT 14
US-09-878-574-14495/C
; Sequence 14495, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 14495
; LENGTH: 163
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701068850H1
US-09-878-574-14495
```

```

Query Match          11.8%; Score 21.8; DB 10; Length 163;
Best Local Similarity 54.3%; Pred. No. 5.1e+02;
Matches 44; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
```

```
OY      9  CTTACGTAACAGACAGTGTGTCCTGTTGCTACTTAAAGACGGTGT 68
          | | | | | | | | | | | | | | | | | | | | | | | |
DB     160 CGTTACTATAGCAAGGTTTGTCTTCTGTTGTTGAAGCTGTGGCGTGGGAA 101
OY      69  CTCGAATTGATTAAGACCATG 89
          | | | | | | | | | | | | | | | | | | | | | | | |
DB     100 TTGAGAGTGAAGAAAGGCATG 80
```

RESULT 15

```
US-10-137-765-52/c
; Sequence 52, Application US/10137765
; Publication No. US20030028926A1
; GENERAL INFORMATION:
; APPLICANT: KUMAGAI, Monto H.
; APPLICANT: DELLA-CIOPPA, Guy R.
; APPLICANT: ERWIN, Robert L.
; APPLICANT: MGEE, David R.
; TITLE OF INVENTION: METHOD OF COMPILING A FUNCTIONAL GENE PROFILE BY
; TITLE OF INVENTION: TRANSDUCTION OF A NON-PLANT DONOR INTO A HOST PLANT IN A POSITIVE
; TITLE OF INVENTION: ACID SEQUENCE OF A NON-PLANT DONOR INTO A HOST PLANT IN A POSITIVE
; FILE REFERENCE: 0080101370507
; CURRENT APPLICATION NUMBER: US/10/137,765
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US/09/359,300A
; PRIOR FILING DATE: 2002-04-29
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: fastseq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 167
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-137-765-52
```

Query Match

11.8%; Score 21.8; DB 9; Length 167;

Best Local Similarity 54.3%; Pred. No. 5.2e+02;

Matches 44; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
OY      77  TGAATAGACCAATGGCAAAACCTGTGACGATTTGGCTTGGATACCCCTGTGCGGTT 136
          | | | | | | | | | | | | | | | | | | | | | | | |
DB     85  TTATAAGACGATTTTGTACTGCTGTGTTGTTGGATGAGCAACAAACATGAGTTT 26
OY     137  CTATTGCACAGAAATCAGAC 157
          | | | | | | | | | | | | | | | | | | | | | | | |
DB      25  GGTTTAGTATGAACAGATC 5
```

Search completed: February 16, 2003, 09:05:11
Job time : 100 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 06:16:28 : Search time 3790 Seconds
(without alignments)
1220.638 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239
Perfect score: 184
Sequence: 1 actcattcttcaggtacag.....ccttagtaggaagcattga 184

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 15899822

Minimum DB seq length: 0
Maximum DB seq length: 184

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

- 1: /cgn2_6/ptodata/2/pna/PCYUS.COMB.seq:*
- 2: /cgn2_6/ptodata/2/pna/US06.COMB.seq:*
- 3: /cgn2_6/ptodata/2/pna/US07.COMB.seq:*
- 4: /cgn2_6/ptodata/2/pna/US080.COMB.seq:*
- 5: /cgn2_6/ptodata/2/pna/US081.COMB.seq:*
- 6: /cgn2_6/ptodata/2/pna/US082.COMB.seq:*
- 7: /cgn2_6/ptodata/2/pna/US083.COMB.seq:*
- 8: /cgn2_6/ptodata/2/pna/US084.COMB.seq:*
- 9: /cgn2_6/ptodata/2/pna/US085.COMB.seq:*
- 10: /cgn2_6/ptodata/2/pna/US086.COMB.seq:*
- 11: /cgn2_6/ptodata/2/pna/US087.COMB.seq:*
- 12: /cgn2_6/ptodata/2/pna/US088.COMB.seq:*
- 13: /cgn2_6/ptodata/2/pna/US089.COMB.seq:*
- 14: /cgn2_6/ptodata/2/pna/US090.COMB.seq:*
- 15: /cgn2_6/ptodata/2/pna/US091.COMB.seq:*
- 16: /cgn2_6/ptodata/2/pna/US092.COMB.seq:*
- 17: /cgn2_6/ptodata/2/pna/US093.COMB.seq:*
- 18: /cgn2_6/ptodata/2/pna/US094.COMB.seq:*
- 19: /cgn2_6/ptodata/2/pna/US095A.COMB.seq:*
- 20: /cgn2_6/ptodata/2/pna/US095B.COMB.seq:*
- 21: /cgn2_6/ptodata/2/pna/US095C.COMB.seq:*
- 22: /cgn2_6/ptodata/2/pna/US095D.COMB.seq:*
- 23: /cgn2_6/ptodata/2/pna/US096A.COMB.seq:*
- 24: /cgn2_6/ptodata/2/pna/US096B.COMB.seq:*
- 25: /cgn2_6/ptodata/2/pna/US096C.COMB.seq:*
- 26: /cgn2_6/ptodata/2/pna/US096D.COMB.seq:*
- 27: /cgn2_6/ptodata/2/pna/US096E.COMB.seq:*
- 28: /cgn2_6/ptodata/2/pna/US097A.COMB.seq:*
- 29: /cgn2_6/ptodata/2/pna/US097B.COMB.seq:*
- 30: /cgn2_6/ptodata/2/pna/US097C.COMB.seq:*
- 31: /cgn2_6/ptodata/2/pna/US098A.COMB.seq:*
- 32: /cgn2_6/ptodata/2/pna/US098B.COMB.seq:*
- 33: /cgn2_6/ptodata/2/pna/US098C.COMB.seq:*
- 34: /cgn2_6/ptodata/2/pna/US099A.COMB.seq:*
- 35: /cgn2_6/ptodata/2/pna/US099B.COMB.seq:*
- 36: /cgn2_6/ptodata/2/pna/US099C.COMB.seq:*
- 37: /cgn2_6/ptodata/2/pna/US099D.COMB.seq:*
- 38: /cgn2_6/ptodata/2/pna/US100A.COMB.seq:*
- 39: /cgn2_6/ptodata/2/pna/US100B.COMB.seq:*
- 40: /cgn2_6/ptodata/2/pna/US101A.COMB.seq:*
- 41: /cgn2_6/ptodata/2/pna/US101B.COMB.seq:*
- 42: /cgn2_6/ptodata/2/pna/US102A.COMB.seq:*
- 43: /cgn2_6/ptodata/2/pna/US102B.COMB.seq:*

| Result | | | | | Query | | | | | SUMMARIES | | | | | Description | | | | |
|--------|-------|-------|--------|----|----------------------|-------|-------|--------|-----|-----------|-------|-------|--------|-----|-------------|-------|-------|--------|----|
| No. | Score | Match | Length | ID | No. | Score | Match | Length | ID | No. | Score | Match | Length | ID | No. | Score | Match | Length | ID |
| 1 | 79.8 | 43.0 | 91 | 17 | US-09-396-087-2094 | 44: | 184 | 100 | 184 | 44: | 184 | 100 | 184 | 184 | 1 | 79.8 | 43.0 | 91 | 17 |
| 2 | 73.6 | 40.0 | 145 | 53 | US-09-540-213-47680 | 45: | 184 | 100 | 184 | 45: | 184 | 100 | 184 | 184 | 2 | 73.6 | 40.0 | 145 | 53 |
| 3 | 73.6 | 40.0 | 145 | 53 | US-09-540-213-47680 | 46: | 184 | 100 | 184 | 46: | 184 | 100 | 184 | 184 | 3 | 73.6 | 40.0 | 145 | 53 |
| 4 | 62.6 | 34.0 | 143 | 69 | US-60-253-653-2055 | 47: | 184 | 100 | 184 | 47: | 184 | 100 | 184 | 184 | 4 | 62.6 | 34.0 | 143 | 69 |
| 5 | 57 | 31.0 | 113 | 15 | US-09-113-753-287 | 48: | 184 | 100 | 184 | 48: | 184 | 100 | 184 | 184 | 5 | 57 | 31.0 | 113 | 15 |
| 6 | 57 | 31.0 | 113 | 21 | US-09-540-212A-64473 | 49: | 184 | 100 | 184 | 49: | 184 | 100 | 184 | 184 | 6 | 57 | 31.0 | 113 | 21 |
| 7 | 57 | 31.0 | 113 | 21 | US-09-540-212A-64473 | 50: | 184 | 100 | 184 | 50: | 184 | 100 | 184 | 184 | 7 | 57 | 31.0 | 113 | 21 |
| 8 | 51 | 27.7 | 51 | 76 | US-60-324-185-16164 | 51: | 184 | 100 | 184 | 51: | 184 | 100 | 184 | 184 | 8 | 51 | 27.7 | 51 | 76 |
| 9 | 30.4 | 16.5 | 172 | 25 | US-09-654-617-200647 | 52: | 184 | 100 | 184 | 52: | 184 | 100 | 184 | 184 | 9 | 30.4 | 16.5 | 172 | 25 |
| 10 | 30.4 | 16.5 | 172 | 27 | US-09-684-016-200647 | 53: | 184 | 100 | 184 | 53: | 184 | 100 | 184 | 184 | 10 | 30.4 | 16.5 | 172 | 27 |
| 11 | 27.6 | 15.0 | 120 | 25 | US-09-644-867-1868 | 54: | 184 | 100 | 184 | 54: | 184 | 100 | 184 | 184 | 11 | 27.6 | 15.0 | 120 | 25 |
| 12 | 27 | 14.7 | 148 | 13 | US-08-940-864-2267 | 55: | 184 | 100 | 184 | 55: | 184 | 100 | 184 | 184 | 12 | 27 | 14.7 | 148 | 13 |
| 13 | 27 | 14.7 | 148 | 20 | US-09-539-806-3010 | 56: | 184 | 100 | 184 | 56: | 184 | 100 | 184 | 184 | 13 | 27 | 14.7 | 148 | 20 |
| 14 | 27 | 14.7 | 148 | 20 | US-09-539-806-3010 | 57: | 184 | 100 | 184 | 57: | 184 | 100 | 184 | 184 | 14 | 27 | 14.7 | 148 | 20 |
| 15 | 27 | 14.7 | 148 | 46 | US-60-027-236-2267 | 58: | 184 | 100 | 184 | 58: | 184 | 100 | 184 | 184 | 15 | 27 | 14.7 | 148 | 46 |
| 16 | 26.6 | 14.5 | 167 | 13 | US-10-929-386-22132 | 59: | 184 | 100 | 184 | 59: | 184 | 100 | 184 | 184 | 16 | 26.6 | 14.5 | 167 | 13 |
| 17 | 26.6 | 14.5 | 175 | 13 | US-08-923-902-3701 | 60: | 184 | 100 | 184 | 60: | 184 | 100 | 184 | 184 | 17 | 26.6 | 14.5 | 175 | 13 |
| 18 | 26.6 | 14.5 | 175 | 21 | US-09-540-208-32804 | 61: | 184 | 100 | 184 | 61: | 184 | 100 | 184 | 184 | 18 | 26.6 | 14.5 | 175 | 21 |
| 19 | 26.2 | 14.2 | 157 | 17 | US-09-306-349-44559 | 62: | 184 | 100 | 184 | 62: | 184 | 100 | 184 | 184 | 19 | 26.2 | 14.2 | 157 | 17 |
| 20 | 26.2 | 14.2 | 157 | 36 | US-09-960-481-44559 | 63: | 184 | 100 | 184 | 63: | 184 | 100 | 184 | 184 | 20 | 26.2 | 14.2 | 157 | 36 |
| 21 | 25.6 | 13.9 | 162 | 1 | PCT-US01-00663-21878 | 64: | 184 | 100 | 184 | 64: | 184 | 100 | 184 | 184 | 21 | 25.6 | 13.9 | 162 | 1 |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

```
22 25.6 13.9 162 33 US-09-864-761-23730 Sequence 23730, A
23 25.6 13.9 162 41 US-10-182-993-21190 Sequence 21190, A
24 25.6 13.9 162 41 US-10-182-993-16876 Sequence 16876, A
25 25.6 13.9 162 41 US-10-182-997-15850 Sequence 15850, A
26 25.6 13.9 162 42 US-10-203-134-21913 Sequence 21913, A
27 25.6 13.9 162 42 US-10-203-135-21502 Sequence 21502, A
28 25.6 13.9 162 42 US-10-203-136-22090 Sequence 22090, A
29 25.6 13.9 162 42 US-10-203-137-21878 Sequence 21878, A
30 25.6 13.9 162 42 US-10-203-139-21075 Sequence 21075, A
31 25.4 13.8 154 63 US-09-849-529A-23897 Sequence 23897, A
32 25.4 13.8 154 63 US-09-849-529A-23897 Sequence 23897, A
33 25.4 13.8 158 31 US-09-823-241-6658 Sequence 23875, A
34 25.2 13.7 150 12 US-08-842-978-716 Sequence 6658, A
35 25.2 13.7 150 12 US-08-842-978-716 Sequence 716, App
36 25.2 13.6 109 18 US-09-436-762A-25245 Sequence 25245, A
37 25.2 13.6 110 18 US-09-436-762A-25245 Sequence 32976, A
38 25.2 13.6 130 25 US-09-652-355-1590 Sequence 1590, App
39 25.2 13.6 148 15 US-09-169-662-1883 Sequence 1883, App
40 24.8 13.5 146 13 US-08-829-307-1926 Sequence 1926, App
41 24.8 13.5 146 21 US-09-540-229-43380 Sequence 43380, A
42 24.8 13.5 146 47 US-60-031-419-1213 Sequence 1213, App
43 24.8 13.5 148 16 US-60-032-577-1213 Sequence 1213, App
44 24.8 13.5 148 16 US-09-294-093A-909 Sequence 909, App
45 24.8 13.5 148 16 US-09-294-093B-803 Sequence 803, App
```

ALIGNMENTS

```
RESULT 1
US-09-396-087-2094
; Sequence 2094, Application US/09396087
; GENERAL INFORMATION:
; APPLICANT: Gearling, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: PULMONARY ARTERY SMOOTH MUSCLE LIBRARY
; FILE REFERENCE: M198-39PM
; CURRENT APPLICATION NUMBER: US/09/396,087
; EARLIER FILING DATE: 1999-09-14
; EARLIER APPLICATION NUMBER: 60/100,260
; EARLIER FILING DATE: 1998-09-14
; EARLIER APPLICATION NUMBER: 60/107,226
; EARLIER FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: 60/131,810
; EARLIER FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5220
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2094
; LENGTH: 91
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(91)
; OTHER INFORMATION: n = A,T,C or G
US-09-396-087-2094
```

```
Query Match 43.4%; Score 79.8; DB 17; Length 91;
Best Local Similarity 93.1%; Pred. No. 3e-15;
Matches 81; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 ACTTCATCTTCAGGTACAGACAGTGTGTGCTGCTGCTGCTGCTAAGGCTTAAGAG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 ACTTCATCTTCAGGTACAGACAGTGTGTGCTGCTGCTGCTGCTAAGGCTTAAGAG 64
QY 61 CGGTGTTCTCAGATTAAGACCA 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 CGGTGTTCTCAGATTAAGACCA 91
```

```
RESULT 2
US-09-540-213-47680
```

```
; Sequence 47680, Application US/09540213
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleage, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS
; FILE REFERENCE: PD-1031 CIP
; CURRENT APPLICATION NUMBER: US/09/540,213
; EARLIER FILING DATE: 2000-03-31
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 53844
; SEQ ID NO 47680
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: rat00177585
US-09-540-213-47680
```

```
Query Match 40.0%; Score 73.6; DB 21; Length 145;
Best Local Similarity 75.8%; Pred. No. 3.6e-13;
Matches 91; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
```

```
QY 9 CTTGAGTACAGACAGTGTGTGCTGCTGCTGCTGCTGCTTAAGAGCGGTGTT 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 CTTGCTGTTGAGGTACAGACAGCTTTCTGCTGCTGCTGCTGCTTAAGAGCGGTGTT 84
QY 69 CTCGAAATTGATTAAGACCATGGCACAACCTGTGACGATGCTTTGAGTAACCTGTG 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 CTCGAAATTGATTAAGACCATGGCACAACCTATGTGACGATGCTTTGAGTAACCTGTG 144
```

```
RESULT 3
US-60-096-664-2067
; Sequence 2067, Application US/60096664
; GENERAL INFORMATION:
; APPLICANT: Stuve, Laura L.
; APPLICANT: Goralski, Thomas J.
; APPLICANT: Garrow, Bonnie L.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM TREATED RAT LIVER
; FILE REFERENCE: P2-0092 P
; CURRENT APPLICATION NUMBER: US/60/096,664
; EARLIER FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 2071
; SOFTWARE: PERL Program
; SEQ ID NO 2067
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE: -
; OTHER INFORMATION: 701342469H1
US-60-096-664-2067
```

```
Query Match 40.0%; Score 73.6; DB 53; Length 145;
Best Local Similarity 75.8%; Pred. No. 3.6e-13;
Matches 91; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
```

```
QY 9 CTTGAGTACAGACAGTGTGTGCTGCTGCTGCTGCTGCTTAAGAGCGGTGTT 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 CTTGCTGTTGAGGTACAGACAGCTTTCTGCTGCTGCTGCTGCTTAAGAGCGGTGTT 84
QY 69 CTCGAAATTGATTAAGACCATGGCACAACCTGTGACGATGCTTTGAGTAACCTGTG 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 CTCGAAATTGATTAAGACCATGGCACAACCTATGTGACGATGCTTTGAGTAACCTGTG 144
```

```
RESULT 4
US-60-253-20525
; Sequence 20525, Application US/60253653
```

```
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions isolated from ovine tissues
; FILE REFERENCE: 1057P
; CURRENT APPLICATION NUMBER: US/60/253,653
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 30124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 20525
; LENGTH: 143
; TYPE: DNA
; ORGANISM: Ovine
; US-60-253-653-20525

Query Match          34.0%; Score 62.6; DB 69; Length 143;
Best Local Similarity 79.6%; Pred. No. 1.3e-09;
Matches 74; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 6 ATCTCAGTACGACAGGTGTTGTGCTTCTGCTGCTGCTTACCTTAGACGCGTG 65
Db 51 ATCTCGCTTCCTCCAGTACGACAGCATGTTCAGTGTGCTAGCTTAGAAGCGCTG 110

QY 66 TTCTCAGATTGATTAGACCATGGCACAAAC 98
Db 111 TTCTCGAATTGATTAGACCATGGCACAAAC 143

RESULT 5
US-09-113-753-287
; Sequence 287, Application US/09113753
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM RAT
; TITLE OF INVENTION: LUNG
; NUMBER OF SEQUENCES: 1178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,753
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/052,200
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PZ-0006P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 287:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
```

```
; CLONE: 700178734H1
US-09-113-753-287

Query Match          31.0%; Score 57; DB 15; Length 113;
Best Local Similarity 77.5%; Pred. No. 8e-08;
Matches 69; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 9 CTCAGGTACAGACAGTGTGCTGCTGCTGCTGCTTACCTTAGAAGCGGTGTT 68
Db 25 CTTCCTGTTCAGTACGACAGCGTTTCTGCTGCTGCTAGCTTAGAAGCGGTGTT 84

QY 69 CTCGAATTGATTAGACCATGGCACAAAA 97
Db 85 CTCGAATTGATTAGACCATGGCACAAAA 113

RESULT 6
US-09-540-212A-64473
; Sequence 64473, Application US/09540212A
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mollaby, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE
; FILE REFERENCE: PD-1034 CIP
; CURRENT APPLICATION NUMBER: US/09/540,212A
; CURRENT FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 67551
; SOFTWARE: PERL Program
; SEQ ID NO: 64473
; LENGTH: 113
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: rat00232595
; US-09-540-212A-64473

Query Match          31.0%; Score 57; DB 21; Length 113;
Best Local Similarity 77.5%; Pred. No. 8e-08;
Matches 69; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 9 CTCAGGTACAGACAGTGTGCTGCTGCTGCTGCTTACCTTAGAAGCGGTGTT 68
Db 25 CTTCCTGTTCAGTACGACAGCGTTTCTGCTGCTGCTAGCTTAGAAGCGGTGTT 84

QY 69 CTCGAATTGATTAGACCATGGCACAAAA 97
Db 85 CTCGAATTGATTAGACCATGGCACAAAA 113

RESULT 7
US-60-052-200-287
; Sequence 287, Application US/60052200
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM RAT
; TITLE OF INVENTION: LUNG
; NUMBER OF SEQUENCES: 1178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
```

QY 89 GGCACAAACTGTGACGATTGGCTTTGGAGTAACCTGTGTGGCGTTCCTA 139
|||||
Db 1 GGCACAAACTGTGACGATTGGCTTTGGAGTAACCTGTGTGGCGTTCCTA 51

```

RESULT 11
US-09-644-867-1868/C
; Sequence 1868, Application US/09644867
GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Donovan, Michael J.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Culpepper, Janice A.
TITLE OF INVENTION: NOVEL NOCTEMIC ACID MOLECULES AND USES

```



```
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1173-001
; CURRENT APPLICATION NUMBER: US/09/644,867
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,061
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 8090
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1868
; LENGTH: 120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-644-867-1868

Query Match          15.0%; Score 27.6; DB 25; Length 120;
Best Local Similarity 60.8%; Pred. No. 2.9e+02;
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

OY 25 TGTGTGTCGTCCTGTCGCTACTTAAAGACGGTGTTCACGAATGATAGA 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 TTTTTCATTTTTCATTAATGTGACAAATATATTTAGATATATTTAAAGATTTAAAAA 55

OY 85 CCATGCGACAAAC 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54 ACATTTCACAAAC 41

RESULT 12
US-08-940-864-2267/c
; Sequence 2267, Application US/08940864
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuv, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Ted M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN BREAST
; NUMBER OF SEQUENCES: 3272
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,864
; FILING DATE: HEREWITH
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,236
; FILING DATE: SEPTEMBER 30, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CERROWE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0240 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 2267:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 1921229
US-08-940-864-2267

Query Match          14.7%; Score 27; DB 13; Length 148;
Best Local Similarity 55.2%; Pred. No. 4.8e+02;
Matches 48; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY 10 TTCAGTACAGACAGTGTGTGTCGTCGCTACTTAAAGACGGGTTC 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 TACANGTAAGANNAGTGAAGTTGTCATTGGGGGNTCCATAAGGGCCTCATGTG 74

OY 70 TCAGAAATTGATAAGACCAGTGCACAAA 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 TCTGAAGGGGTACACCTTAGGAAACA 47

RESULT 13
US-09-539-806-3010/c
; Sequence 3010, Application US/09539806
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE
; FILE REFERENCE: PD-1027 CIP
; CURRENT APPLICATION NUMBER: US/09/539,806
; CURRENT FILING DATE: 2000-03-30
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 48372
; SOFTWARE: PERL Program
; SEQ ID NO 3010
; LENGTH: 148
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Inocyte ID No: hu00498291
; NAME/KEY: unsure
; LOCATION: 27, 30-31, 38, 40, 95, 120-121, 129, 140
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-806-3010

Query Match          14.7%; Score 27; DB 20; Length 148;
Best Local Similarity 55.2%; Pred. No. 4.8e+02;
Matches 48; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY 10 TTCAGTACAGACAGTGTGTGTCGTCGCTACTTAAAGACGGGTTC 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 TACANGTAAGANNAGTGAAGTTGTCATTGGGGGNTCCATAAGGGCCTCATGTG 74

OY 70 TCAGAAATTGATAAGACCAGTGCACAAA 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 TCTGAAGGGGTACACCTTAGGAAACA 47

RESULT 14
US-09-539-806B-3010/c
; Sequence 3010, Application US/09539806B
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF EXOCRINE GLAND TISSUE
```

FILE REFERENCE: PD-1027 CIP
CURRENT APPLICATION NUMBER: US/09/539, 806B
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 08/706,766
PRIOR FILING DATE: September 27, 1996
PRIOR APPLICATION NUMBER: 60/004,676
PRIOR FILING DATE: October 2, 1995
PRIOR APPLICATION NUMBER: 08/749,515
PRIOR FILING DATE: November 15, 1996
PRIOR APPLICATION NUMBER: 60/006,810
PRIOR FILING DATE: November 15, 1995
PRIOR APPLICATION NUMBER: 08/822,285
PRIOR FILING DATE: March 20, 1997
PRIOR APPLICATION NUMBER: 60/013,696
PRIOR FILING DATE: March 20, 1996
PRIOR APPLICATION NUMBER: 08/951,197
PRIOR FILING DATE: October 1, 1997
PRIOR APPLICATION NUMBER: 60/027,249
PRIOR FILING DATE: October 1, 1996
PRIOR APPLICATION NUMBER: 08/826,438
PRIOR FILING DATE: March 20, 1997
PRIOR APPLICATION NUMBER: 60/016,145
PRIOR FILING DATE: April 18, 1996
PRIOR APPLICATION NUMBER: 60/013,696
PRIOR FILING DATE: March 20, 1996
PRIOR APPLICATION NUMBER: 08/839,389
PRIOR FILING DATE: April 11, 1997
PRIOR APPLICATION NUMBER: 60/015,312
PRIOR FILING DATE: April 12, 1996
PRIOR APPLICATION NUMBER: 08/951,198
PRIOR FILING DATE: October 1, 1997
PRIOR APPLICATION NUMBER: 60/027,249
PRIOR FILING DATE: October 1, 1996
PRIOR APPLICATION NUMBER: 08/839,968
PRIOR FILING DATE: April 24, 1997
PRIOR APPLICATION NUMBER: 60/016,150
PRIOR FILING DATE: April 24, 1996
PRIOR APPLICATION NUMBER: 08/846,104
PRIOR FILING DATE: April 25, 1997
PRIOR APPLICATION NUMBER: 60/016,838
PRIOR FILING DATE: April 26, 1996
PRIOR APPLICATION NUMBER: 08/999,886
PRIOR FILING DATE: May 13, 1997
PRIOR APPLICATION NUMBER: 60/017,555
PRIOR FILING DATE: May 13, 1996
PRIOR APPLICATION NUMBER: 60/031,667
PRIOR FILING DATE: November 11, 1996
PRIOR APPLICATION NUMBER: 08/901,902
PRIOR FILING DATE: July 24, 1997
PRIOR APPLICATION NUMBER: 60/023,379
PRIOR FILING DATE: July 26, 1996
PRIOR APPLICATION NUMBER: 08/918,181
PRIOR FILING DATE: August 27, 1997
PRIOR APPLICATION NUMBER: 60/023,236
PRIOR FILING DATE: August 28, 1996
PRIOR APPLICATION NUMBER: 08/940,864
PRIOR FILING DATE: September 29, 1997
PRIOR APPLICATION NUMBER: 60/027,236
PRIOR FILING DATE: September 30, 1996
PRIOR APPLICATION NUMBER: 08/956,502
PRIOR FILING DATE: October 22, 1997
PRIOR APPLICATION NUMBER: 60/029,083
PRIOR FILING DATE: October 23, 1996
PRIOR APPLICATION NUMBER: 08/993,402
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/033,647
PRIOR FILING DATE: December 19, 1996
PRIOR APPLICATION NUMBER: 09/022,355
PRIOR FILING DATE: February 11, 1998
PRIOR APPLICATION NUMBER: 60/044,847
PRIOR FILING DATE: February 13, 1997
PRIOR APPLICATION NUMBER: 09/045,574
PRIOR FILING DATE: March 20, 1998

PRIOR APPLICATION NUMBER: 60/041,275
PRIOR FILING DATE: March 21, 1997
PRIOR APPLICATION NUMBER: 09/042,629
PRIOR FILING DATE: March 16, 1998
PRIOR APPLICATION NUMBER: 60/043,613
PRIOR FILING DATE: March 18, 1997
PRIOR APPLICATION NUMBER: 09/057,988
PRIOR FILING DATE: April 9, 1998
PRIOR APPLICATION NUMBER: 60/043,256
PRIOR FILING DATE: April 16, 1997
PRIOR APPLICATION NUMBER: 09/114,041
PRIOR FILING DATE: July 10, 1998
PRIOR APPLICATION NUMBER: 60/052,257
PRIOR FILING DATE: July 10, 1997
PRIOR APPLICATION NUMBER: 09/145,340
PRIOR FILING DATE: September 1, 1998
PRIOR APPLICATION NUMBER: 60/058,922
PRIOR FILING DATE: September 11, 1997
NUMBER OF SEQ ID NOS: 48372
SOFTWARE: PERL Program
SEQ ID NO 3010
LENGTH: 148
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: hu00498291
FEATURE:
NAME/KEY: unsure
LOCATION: 27, 30-31, 38, 40, 95, 120-121, 129, 140
OTHER INFORMATION: a, t, c, g, or other
US-09-539-806B-3010

Query Match 14.7% Score 27: DB 20: Length 148:
Best Local Similarity 55.2% Pred. No. 4.8e+02:
Matches 48: Conservative 0: Mismatches 39: Indels 0: Gaps 0:

OY 10 TTCAGTACAGACAGTGTGTTGTCCTGCTGCTTAAGAGCGGTGTTTC 69
DB 133 TACAGCTAAGAGNAGAGTGTTCATTTGGGGGNTCCATNAGGCGCATGTG 74
OY 70 TCAGATTGATTAAGACATGCGACAAA 96
DB 73 TCTGAAGGGTAACACCTAGGGAACA 47

RESULT 15
US-60-027-236-2267/c
Sequence 2267, Application US/60027236
GENERAL INFORMATION:
APPLICANT: Gooding, Douglas H.
APPLICANT: Stuve, Laura L.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Akedblom, Ingrid E.
APPLICANT: Deleage, Angelo M.
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
NUMBER OF SEQUENCES: 3272
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/60/027,236
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: CERRONE, MICHAEL C., Ph.D.
: REGISTRATION NUMBER: 39132
: REFERENCE/DOCKET NUMBER: PD-0240P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 855-0555
: TELEFAX: (415) 845-4166
: INFORMATION FOR SEQ ID NO: 2267:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 148 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: 1921229
: US-60-027-236-2267

```

```

Query Match          14.7%: Score 27; DB 46; Length 148;
Best Local Similarity 55.2%: Pred. No. 4.8e+02;
Matches 48; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY 10 TTCAGGTACAGACAGTGTGTTGCTGTTGCTGCTACTTTAAGAAGCGGTGTTTC 69
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 133 TACANGTAAGANNAAGTGAAGTTGTTCCATTTGGGGGNTCCATAAGGGGCTCATGTG 74
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 70 TCAGAAATTGATAGACCATGGCACAAA 96
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 73 TCTGAAGGGGTACACCTAAGGAACA 47
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: February 16, 2003, 08:56:57
Job time : 3792 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 06:21:42 : Search time 368 Seconds
(without alignments)
745.158 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239

Perfect score: 184
Sequence: 1 actcatctcttcaggtacag.....cctagtagtaagatgta 184

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2773584 seqs, 745158349 residues

Total number of hits satisfying chosen parameters: 4672920

Minimum DB seq length: 0
Maximum DB seq length: 184

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| C 1 | 25 | 13.6 | 148 | US-09-534-850-23238 | Sequence 23238, A |
| C 2 | 25 | 13.6 | 158 | US-09-513-999C-34505 | Sequence 34505, A |
| C 3 | 24.8 | 13.5 | 181 | US-09-534-850-7252 | Sequence 7252, Ap |
| C 4 | 23.6 | 12.8 | 169 | US-09-513-999C-35609 | Sequence 35609, A |
| C 5 | 23.4 | 12.7 | 137 | US-10-314-344-17067 | Sequence 17067, A |
| C 6 | 22.8 | 12.4 | 157 | US-09-513-999C-12241 | Sequence 12241, A |
| C 7 | 22.4 | 12.2 | 132 | US-10-286-136-975 | Sequence 975, App |
| C 8 | 22.4 | 12.2 | 145 | US-09-531-113-20439 | Sequence 20439, A |
| C 9 | 22.2 | 12.1 | 161 | US-09-606-397A-276 | Sequence 276, App |
| C 10 | 22.2 | 12.1 | 172 | US-09-513-999C-20707 | Sequence 20707, A |
| C 11 | 22 | 12.0 | 149 | US-10-314-344-28480 | Sequence 28480, A |
| C 12 | 21.6 | 11.7 | 154 | US-09-620-607B-198 | Sequence 198, App |
| C 13 | 21.6 | 11.7 | 170 | US-10-314-344-28592 | Sequence 28592, A |
| C 14 | 21.4 | 11.6 | 126 | US-10-329-624-1273 | Sequence 1273, Ap |
| C 15 | 21.4 | 11.6 | 143 | US-10-286-136-913 | Sequence 913, App |
| C 16 | 21.2 | 11.5 | 130 | US-10-335-066-1298 | Sequence 1298, Ap |
| C 17 | 21.2 | 11.5 | 121 | US-09-531-113-2842 | Sequence 2842, Ap |
| C 18 | 21 | 11.4 | 145 | US-10-314-344-8580 | Sequence 8580, Ap |
| C 19 | 21 | 11.4 | 175 | US-10-203-138A-6831 | Sequence 6831, Ap |
| C 20 | 21 | 11.4 | 180 | US-09-531-113-36300 | Sequence 36300, A |
| C 21 | 21 | 11.4 | 183 | US-09-513-999C-9763 | Sequence 9763, Ap |
| C 22 | 20.8 | 11.3 | 115 | US-09-513-999C-14614 | Sequence 14614, A |
| C 23 | 20.8 | 11.3 | 148 | US-10-286-136-567 | Sequence 567, App |
| C 24 | 20.8 | 11.3 | 160 | US-09-513-999C-20427 | Sequence 20427, A |
| C 25 | 20.8 | 11.3 | 160 | US-09-513-999C-22766 | Sequence 22766, A |
| C 26 | 20.8 | 11.3 | 168 | US-10-144-771-46425 | Sequence 46425, A |

| | | | | | | |
|------|------|------|-----|---|----------------------|--------------------|
| 27 | 20.8 | 11.3 | 184 | 5 | US-09-513-999C-16606 | Sequence 16606, A |
| 28 | 20.6 | 11.2 | 89 | 5 | US-10-298-796-38 | Sequence 38, App1 |
| C 29 | 20.6 | 11.2 | 146 | 5 | US-09-534-850-13902 | Sequence 13902, A |
| C 30 | 20.6 | 11.2 | 159 | 5 | US-09-513-999C-36652 | Sequence 36652, A |
| C 31 | 20.6 | 11.2 | 170 | 6 | US-10-266-131-488 | Sequence 488, App |
| C 32 | 20.6 | 11.2 | 183 | 5 | US-09-513-999C-10927 | Sequence 10927, A |
| C 33 | 20.4 | 11.1 | 84 | 5 | US-09-531-113-18049 | Sequence 18049, A |
| C 34 | 20.4 | 11.1 | 103 | 6 | US-10-203-138A-8344 | Sequence 8344, Ap |
| C 35 | 20.4 | 11.1 | 132 | 5 | US-09-513-999C-16924 | Sequence 16924, A |
| C 36 | 20.4 | 11.1 | 146 | 5 | US-09-513-999C-20770 | Sequence 20770, A |
| C 37 | 20.4 | 11.1 | 146 | 5 | US-09-513-999C-23650 | Sequence 23650, A |
| C 38 | 20.4 | 11.1 | 153 | 5 | US-09-513-999C-25375 | Sequence 25375, A |
| C 39 | 20.4 | 11.1 | 179 | 5 | US-09-513-999C-8824 | Sequence 8824, Ap |
| C 40 | 20.2 | 11.0 | 50 | 6 | US-10-329-624-5149 | Sequence 5149, Ap |
| C 41 | 20.2 | 11.0 | 96 | 6 | US-10-314-344-35529 | Sequence 35529, Ap |
| C 42 | 20.2 | 11.0 | 101 | 6 | US-10-314-344-35568 | Sequence 35568, A |
| C 43 | 20.2 | 11.0 | 109 | 6 | US-10-157-169A-50 | Sequence 50, App1 |
| C 44 | 20.2 | 11.0 | 109 | 6 | US-10-157-491A-50 | Sequence 50, App1 |
| C 45 | 20.2 | 11.0 | 109 | 6 | US-10-154-951B-50 | Sequence 50, App1 |

ALIGNMENTS

RESULT 1
US-09-534-850-23238/C
Sequence 23238, Application US/09534850
GENERAL INFORMATION:
APPLICANT: Selhamer, Jeffrey J.
APPLICANT: Deleane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING METABOLIC MOLECULES
FILE REFERENCE: PD-1020 CIP
CURRENT APPLICATION NUMBER: US/09/534, 850
CURRENT FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 07/916, 491
PRIOR FILING DATE: 1992-07-17
PRIOR APPLICATION NUMBER: 07/977, 780
PRIOR FILING DATE: 1992-11-19
PRIOR APPLICATION NUMBER: 08/100, 523
PRIOR FILING DATE: 1993-08-03
PRIOR APPLICATION NUMBER: 09/008, 119
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/196, 364
PRIOR FILING DATE: 1994-02-14
PRIOR APPLICATION NUMBER: 08/282, 991
PRIOR FILING DATE: 1994-07-28
PRIOR APPLICATION NUMBER: 08/438, 571
PRIOR FILING DATE: 1995-05-10
PRIOR APPLICATION NUMBER: 08/179, 873
PRIOR FILING DATE: 1994-01-11
PRIOR APPLICATION NUMBER: 08/504, 732
PRIOR FILING DATE: 1995-07-20
PRIOR APPLICATION NUMBER: 08/137, 951
PRIOR FILING DATE: 1993-10-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 26730
SOFTWARE: PERL Program
SEQ ID NO 23238
LENGTH: 148
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: rat00036515
NAME/KEY: unsure
LOCATION: 121, 130, 133, 137-138, 142
OTHER INFORMATION: a, t, c, g, or other
US-09-534-850-23238

```
Query Match 13.6%; Score 25; DB 5; Length 148;
Best Local Similarity 57.1%; Pred. No. 1.1e+02;
Matches 40; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Oy 108 TCGCTTGAGTACCGCTGTGCGTTCCATTTGCACAGAAATCAGAGCCCTATCCCT 167
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141 TGANNCTTNGNAGCCCTTNAACCTGACTCTTGCACAGAGTACAGCCGCTATCCT 82

Oy 168 TAGTAGTGAA 177
    ||| | | | | |
Db 81 TATTAGAGCA 72

RESULT 2
US-09-513-999C-34505/c
; Sequence 34505, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 34505
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 59
; OTHER INFORMATION: k-g or t

US-09-513-999C-34505

Query Match 13.6%; Score 25; DB 5; Length 158;
Best Local Similarity 61.5%; Pred. No. 1.1e+02;
Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Oy 57 GAAGCGGTGTTTCAGATGTGATAGACCATGCAAAACTGTGACGATTGCTTTGG 116
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 158 GAAGGAGCATTCACAGCATTAATAGACACTTGAAAGACAGATTATGCTATGG 99

Oy 117 AGTAA 121
    ||| |
Db 98 AGGAA 94

RESULT 3
US-09-534-850-7252
; Sequence 7252, Application US/09534850
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullaly, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING METABOLIC MOLECULES
; FILE REFERENCE: PD-1020 CIP
; CURRENT APPLICATION NUMBER: US/09/534,850
; CURRENT FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 07/916,491
; PRIOR FILING DATE: 1992-07-17
; PRIOR APPLICATION NUMBER: 07/977,780
; PRIOR FILING DATE: 1992-11-19
; PRIOR APPLICATION NUMBER: 08/100,523
; PRIOR FILING DATE: 1993-08-03
; PRIOR APPLICATION NUMBER: 09/008,119
```

```
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/196,364
; PRIOR FILING DATE: 1994-02-14
; PRIOR APPLICATION NUMBER: 08/282,991
; PRIOR FILING DATE: 1994-07-28
; PRIOR APPLICATION NUMBER: 08/438,571
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: 08/179,873
; PRIOR FILING DATE: 1994-01-11
; PRIOR APPLICATION NUMBER: 08/504,732
; PRIOR FILING DATE: 1995-07-20
; PRIOR APPLICATION NUMBER: 08/137,951
; PRIOR FILING DATE: 1993-10-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 26730
; SOFTWARE: PERL Program
; SEQ ID NO 7252
; LENGTH: 181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00801719
; NAME/KEY: unsure
; LOCATION: 28, 34, 48, 58, 62, 73, 76, 106, 121, 123, 147, 149, 173
; OTHER INFORMATION: a, t, c, g, or other

US-09-534-850-7252

Query Match 13.5%; Score 24.8; DB 5; Length 181;
Best Local Similarity 53.3%; Pred. No. 1.4e+02;
Matches 64; Conservative 0; Mismatches 55; Indels 1; Gaps 1;

Oy 36 TCCGTGTGGCTACTTAAGACGCGTGTTCACAGATGTGATAGACCATGGACAA 95
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 TCCTGTTCCNNTNACTTAATACCTACCTCCACAGACTTTTACACCTGTGCACNT 122

Oy 96 AACTGTGACGATGGCTTTGGAGTAACTGTGCGGTTCTCATTTGACAGAAATCAGA 155
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 NAGCAGAGAGA-TGTCCTGTGACCTTANCNTATGGTCCCTACTAGACCTTTAAACAGA 181

RESULT 4
US-09-513-999C-35609/c
; Sequence 35609, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 35609
; LENGTH: 169
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 136
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 137
; OTHER INFORMATION: s=g or c

US-09-513-999C-35609

Query Match 12.8%; Score 23.6; DB 5; Length 169;
Best Local Similarity 53.4%; Pred. No. 3.3e+02;
```

```
Matches 47; Conservative 1; Mismatches 40; Indels 0; Gaps 0;
OY 52 TTTAAGACGGGTGTTCTCAGATGATAAGACATGGACAAACCTGAGCATGGC 111
    ||||| ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 151 TTTAAAAGAGGCGCTCCAGACAAATGTTCTTTGTCAGTTCTGAGAGTTAACT 92
OY 112 TTTGAGTAACCCCTGTGCGGTTCCCTA 139
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 91 GGTATATAAAGCTGCGCTGCTGCTCGCA 64

RESULT 5
US-10-314-344-17067/C
; Sequence 17067, Application US/10314344
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleage, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING GROWTH, DEVELOPMENT, AND
; TITLE OF INVENTION: DIFFERENTIATION MOLECULES
; FILE REFERENCE: PD-1021-2 CON
; CURRENT APPLICATION NUMBER: US/10/314,344
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: 09/534,846
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 07/916,491
; PRIOR FILING DATE: 1992-07-17
; PRIOR APPLICATION NUMBER: 07/977,780
; PRIOR FILING DATE: 1992-11-19
; PRIOR APPLICATION NUMBER: 08/100,523
; PRIOR FILING DATE: 1993-08-03
; PRIOR APPLICATION NUMBER: 09/008,119
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/196,364
; PRIOR FILING DATE: 1994-02-14
; PRIOR APPLICATION NUMBER: 08/282,991
; PRIOR FILING DATE: 1994-07-28
; PRIOR APPLICATION NUMBER: 08/438,571
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: 08/179,873
; PRIOR FILING DATE: 1994-01-11
; PRIOR APPLICATION NUMBER: 08/504,732
; PRIOR FILING DATE: 1995-07-20
; Remaining prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 38710
; SOFTWARE: PERL Program
; SEQ ID NO 17067
; LENGTH: 137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01250602
US-10-314-344-17067

Query Match 12.7%; Score 23.4; DB 6; Length 137;
Best Local Similarity 52.6%; Pred. No. 3.7e+02;
Matches 51; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
OY 88 TGGCAACAAACCTGACGATTTGGCTTTGGAGTACCCCTGTGGCGTTCTTATTCACAG 147
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 117 TTGGACATATAAGACCTCTTGGAACTAGATTAATGATGTGTAAGTACTGTTATTC 58
OY 148 AATCAGACCTCATTTCCCTTAGTAGTAGTGAACATTGA 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 AAAGCTTAATCATTTACTCTGCTATATGATGAAGA 21

RESULT 6
US-09-513-999C-12241

; Sequence 12241, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent. pm
; SEQ ID NO 12241
; LENGTH: 157
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-12241

Query Match 12.4%; Score 22.8; DB 5; Length 157;
Best Local Similarity 59.1%; Pred. No. 6e+02;
Matches 39; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
OY 39 TGTGTGCTAATTAAGAGCGGTGTTCTCAGATTTGATTAAGACCAACCAAAAC 98
    || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 42 TGCCCTGGCTACTTCTAAAGCCTGTGCTCATTTGCTAATAACAAGAAATGGCTTAAAC 101
OY 99 TGTGAC 104
    || ||
DB 102 TGGAC 107

RESULT 7
US-10-286-136-975
; Sequence 975, Application US/10286136
; GENERAL INFORMATION:
; APPLICANT: Friedlich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Murine Polynucleotide Tags
; TITLE OF INVENTION: and Mutant Cells and Mutant Animals Defined Thereby
; FILE REFERENCE: LEX-0049-USA
; CURRENT APPLICATION NUMBER: US/10/286,136
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US/09/677,367
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US 60/157,364
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 1000
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 975
; LENGTH: 132
; TYPE: DNA
; ORGANISM: mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(132)
; OTHER INFORMATION: n = A,T,C or G
US-10-286-136-975

Query Match 12.2%; Score 22.4; DB 6; Length 132;
Best Local Similarity 49.5%; Pred. No. 7.7e+02;
Matches 50; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
OY 71 CAGAAATGATTAAGACCATGGACCAAAACGTGACGATTTGGAGTAAACCCGTGTGG 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 8 CAGAGATTGAAGACCTNACCGNGTGTATTTCTCTGTTTGGACGAAACCTTTAG 67
OY 131 CGGTTCTATTGACAGAAATCAGACCTTCATTCCTTAGT 171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 68 GGTGACCGTGGCTTNGAGCTCAAAACCGCTGTTTCTTGT 108
```

```
RESULT 8
US-09-531-113-20439/c
; Sequence 20439, Application US/09531113
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 20439
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700955456H1
US-09-531-113-20439

Query Match          12.1%; Score 22.4; DB 5; Length 145;
Best Local Similarity 62.5%; Pred. No. 7.9e+02;
Matches 35; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 4 TCATCTTCAGGTACGACAGTGTGTGTCTCTGTGTGGCTAAGTAA 59
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 88 TCATCTCTTTTTCCTCTTTTCTTTTCTGCTGTGTGACATTAAGTA 33

RESULT 9
US-09-606-397A-276/c
; Sequence 276, Application US/09606397A
; GENERAL INFORMATION:
; APPLICANT: Wattler, Sigrid
; APPLICANT: Nehls, Michael
; APPLICANT: Zandrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Canine Polynucleotides and
; FILE REFERENCE: LEX-0025-USA
; CURRENT APPLICATION NUMBER: US/09/606,397A
; CURRENT FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/142,150
; PRIOR FILING DATE: 1999-07-02
; NUMBER OF SEQ ID NOS: 279
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 276
; LENGTH: 161
; TYPE: DNA
; ORGANISM: Canis Domesticus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 49, 57, 73, 92, 96, 101, 134
; OTHER INFORMATION: n = A,T,C or G
US-09-606-397A-276

Query Match          12.1%; Score 22.2; DB 5; Length 161;
Best Local Similarity 52.5%; Pred. No. 9.4e+02;
Matches 42; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

OY 99 TGTGACGATGTGCTTGTGAGTAACCTGTGCGGTTCCTAATGCACAAATCAGAGCC 158
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 148 TGTGAGGTTGGGTTCAGCATATAAGTCATTTTTCCTCTGTGAGAACNCAGNTCC 89

OY 159 TCATTCCCTAGTAGTAAGA 178
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 88 TGCTTCGCTCGTAGNATG 69

RESULT 10
US-09-513-999C-20707/c
; Sequence 20707, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Ducleert, A.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59,US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 20707
; LENGTH: 172
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-20707

Query Match          12.1%; Score 22.2; DB 5; Length 172;
Best Local Similarity 61.0%; Pred. No. 9.6e+02;
Matches 36; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 24 GTGTTGTGCTGCTCTGTGTGCGCTAAGTAAAGCGGTTTCTCAGAAATGATAA 82
    |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 154 GTGTGTGTGTGTGTGTGTGTGTGTGCGCGCAGTGAACCTCAGATTCTTAAGAAATGATTA 96

RESULT 11
US-10-314-344-28480
; Sequence 28480, Application US/10314344
; GENERAL INFORMATION:
; APPLICANT: Selhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING GROWTH, DEVELOPMENT, AN
; FILE REFERENCE: PD-1021-2 CON
; CURRENT APPLICATION NUMBER: US/10/314,344
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: 09/534,846
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 07/916,491
; PRIOR FILING DATE: 1992-07-17
; PRIOR APPLICATION NUMBER: 07/977,780
; PRIOR FILING DATE: 1992-11-19
; PRIOR APPLICATION NUMBER: 08/100,523
; PRIOR FILING DATE: 1993-08-03
; PRIOR APPLICATION NUMBER: 09/008,119
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/196,364
; PRIOR FILING DATE: 1994-02-14
; PRIOR APPLICATION NUMBER: 08/282,991
; PRIOR FILING DATE: 1994-07-28
; PRIOR APPLICATION NUMBER: 08/438,571
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: 08/179,873
; PRIOR FILING DATE: 1994-01-11
; PRIOR APPLICATION NUMBER: 08/504,732
; PRIOR FILING DATE: 1995-07-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 38710
; SOFTWARE: PERL Program
; SEQ ID NO 28480
; LENGTH: 149
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00938463
; NAME/KEY: unsure
```



```
; LOCATION: 19, 93, 147-148
; OTHER INFORMATION: a, t, c, g, or other
US-10-314-344-28480

Query Match
Best Local Similarity 12.0%; Score 22; DB 6; Length 149;
Matches 49; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY 81 AAGACCATGCGCAAAACGCTGACGATGCGTTTGAGTAACCCCTGCGGCTTCCTAN 140
DB 15 AAGANACTGGAACCAACGCTGCTCCCTACCTCCCTGCACCTACCACTTGAGGATGCGAC 74
OY 141 TGCACAAATACAGACGCTTCATTCCTTAGTAGTGA 176
DB 75 TGAAGAACCTTGAGCGCCNACATACCTCCCTGTGA 110

RESULT 12
US-09-620-607B-198/c
; Sequence 198, Application US/09620607B
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Mutated Mammalian Cells and
; FILE REFERENCE: LEX-0032-USA
; CURRENT APPLICATION NUMBER: US/09/620, 607B
; CURRENT FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/144942
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 1848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 198
; LENGTH: 154
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-620-607B-198

Query Match
Best Local Similarity 11.7%; Score 21.6; DB 5; Length 154;
Matches 33; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 43 GTGGCTACTTTAGAGCGGTGTTCTCGAATTGATAAGCATGCGACA 94
DB 99 GTGGCGCAAGCAATGAGCAGTGTTCTCGAGTGTTATCATACGACAGTCCA 48

RESULT 13
US-10-314-344-28592
; Sequence 28592, Application US/10314344
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING GROWTH, DEVELOPMENT, AND
; FILE REFERENCE: PD-1021-2 CON
; CURRENT APPLICATION NUMBER: US/10/314,344
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: 09/534,846
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 07/916,491
; PRIOR FILING DATE: 1992-07-17
; PRIOR APPLICATION NUMBER: 07/977,780
; PRIOR FILING DATE: 1992-11-19
; PRIOR APPLICATION NUMBER: 08/100,523
; PRIOR FILING DATE: 1993-08-03
; PRIOR APPLICATION NUMBER: 09/008,119
; PRIOR FILING DATE: 1998-01-16
```

```
; PRIOR APPLICATION NUMBER: 08/196,364
; PRIOR FILING DATE: 1994-02-14
; PRIOR APPLICATION NUMBER: 08/282,991
; PRIOR FILING DATE: 1994-07-28
; PRIOR APPLICATION NUMBER: 08/438,571
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: 08/179,873
; PRIOR FILING DATE: 1994-01-11
; PRIOR APPLICATION NUMBER: 08/504,732
; PRIOR FILING DATE: 1995-07-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 38710
; SOFTWARE: PERL Program
; SEQ ID NO: 28592
; LENGTH: 170
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00666057
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 4
; OTHER INFORMATION: a, t, c, g, or other
US-10-314-344-28592

Query Match
Best Local Similarity 11.7%; Score 21.6; DB 6; Length 170;
Matches 54; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

OY 37 CCGTGTGCGCTAACCTTTAGAGCGGTGTTCTCGAATTCATTAAGCATGCGACAA 96
DB 43 CCGCTTCGCTATGTGAGAGGCTCCACTTATGATTCAAAGTCCTACCAAGCGCACTATAA 102
OY 97 ACTGTGACGATGCGCTTGGAGTAACTGCTGTGCGGCTCTATTGCA 144
DB 103 ATTATGAGATTTACTATGATTTATTCATGCTTATATATTGTGAA 150

RESULT 14
US-10-329-624-1273
; Sequence 1273, Application US/10329624
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; APPLICANT: Gil H. Choi
; APPLICANT: Patrick S. Dillon
; APPLICANT: Craig A. Rosen
; APPLICANT: Steven C. Barash
; APPLICANT: Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
```

```
;
; ATTORNEY/AGENT INFORMATION:
;   NAME: Mark J. Hyman
;   REGISTRATION NUMBER: 46,789
;   REFERENCE/DOCKET NUMBER: PB248P1D1
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (240) 314-1224
;     TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 1273:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 126 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1273:
US-10-329-624-1273

Query Match      11.6%; Score 21.4; DB 6; Length 126;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 41 TTGCGCTAAGCTTAAGAGCGGCTTCTCAGATGAT 80
Db 35 TTTGCTTAGCTTAACAATCCGTTTCTCAGACTGAT 74

RESULT 15
US-10-286-136-913
; Sequence 913 Application US/10286136
; GENERAL INFORMATION:
;   APPLICANT: Friedlich, Glenn
;   APPLICANT: Zambrowicz, Brian
;   APPLICANT: Sands, Arthur T.
;   TITLE OF INVENTION: Novel Murine Polynucleotide Tags
;   FILE REFERENCE: LEX-0049-USA
;   CURRENT APPLICATION NUMBER: US/10/286,136
;   PRIOR FILING DATE: 2002-10-31
;   PRIOR APPLICATION NUMBER: US/09/677,367
;   PRIOR FILING DATE: 2000-10-02
;   PRIOR APPLICATION NUMBER: US 60/157,364
;   PRIOR FILING DATE: 1999-10-01
;   NUMBER OF SEQ ID NOS: 1000
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 913
;   LENGTH: 143
;   TYPE: DNA
;   ORGANISM: mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(143)
;   OTHER INFORMATION: n = A,T,C or G
US-10-286-136-913

Query Match      11.6%; Score 21.4; DB 6; Length 143;
Best Local Similarity 51.7%; Pred. No. 1.7e+03;
Matches 46; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 47 CTAAGTTAAGAGCGGCTTCTCAGATGATTAAGACCAACAAACTGTGACGA 106
Db 24 CAAAATATGAACACANNAAGATTCAAGACTCCACAACCTCAATTAATATGACGA 83

OY 107 TTGCGTTGAGTAACCTGTGTGGGTT 135
Db 84 CAGCATAGGANNAACTTGAGTACCTTT 112

Search completed: February 16, 2003, 09:03:18
Job time : 369 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 06:08:23 ; Search time 2200 Seconds
(without alignments)
1354.532 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_239
Perfect score: 184
Sequence: 1 actcattctcagtaacag.....ccttagtagtgaagcattga 184

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 1840244

Minimum DB seq length: 0
Maximum DB seq length: 184

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlun:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_liv:*
20: em_gss_pln:*
21: em_gss_vtc:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 74.6 | 40.5 | 160 | 9 | AA645033 vs86f01.r |
| 2 | 29 | 15.8 | 155 | 10 | AZ071494 RPCI-23-4 |
| 3 | 28.4 | 15.4 | 153 | 10 | BE181187 CM2-HT063 |
| 4 | 27.4 | 14.9 | 157 | 17 | AZ711258 RPCI-24-1 |
| 5 | 26.8 | 14.6 | 152 | 17 | AQ479750 RPCI-11-2 |
| 6 | 26.2 | 14.2 | 179 | 17 | AZ121613 RPCI-23-4 |

| | | | | | |
|----|------|------|-----|----|--------------------|
| 7 | 26 | 14.1 | 142 | 14 | R95649 |
| 8 | 26 | 14.1 | 154 | 12 | BG257473 |
| 9 | 25.8 | 14.0 | 142 | 17 | BH887597 |
| 10 | 25.8 | 14.0 | 178 | 9 | AV281230 |
| 11 | 25.8 | 14.0 | 178 | 14 | BM874696 |
| 12 | 25.6 | 13.9 | 152 | 17 | AZ519722 |
| 13 | 25.4 | 13.8 | 168 | 13 | BG980506 MR3-CM014 |
| 14 | 25.2 | 13.7 | 183 | 10 | BB166886 |
| 15 | 25 | 13.6 | 134 | 17 | AZ230539 |
| 16 | 25 | 13.6 | 154 | 12 | BG361683 |
| 17 | 25 | 13.6 | 183 | 10 | AW781579 |
| 18 | 24.8 | 13.5 | 90 | 9 | AA561369 |
| 19 | 24.6 | 13.4 | 130 | 9 | AA399168 |
| 20 | 24.6 | 13.4 | 150 | 12 | BF251565 |
| 21 | 24.6 | 13.4 | 153 | 9 | AA576737 |
| 22 | 24.6 | 13.4 | 156 | 13 | BG952520 |
| 23 | 24.6 | 13.4 | 159 | 9 | AA464356 |
| 24 | 24.6 | 13.4 | 162 | 9 | AT700924 |
| 25 | 24.6 | 13.4 | 175 | 14 | BQ326519 |
| 26 | 24.6 | 13.4 | 179 | 14 | BQ332549 |
| 27 | 24.6 | 13.4 | 180 | 9 | A1468974 |
| 28 | 24.6 | 13.4 | 180 | 14 | D72192 |
| 29 | 24.6 | 13.4 | 183 | 9 | AA151330 |
| 30 | 24.6 | 13.4 | 183 | 13 | B1050586 |
| 31 | 24.6 | 13.4 | 183 | 13 | B1050592 |
| 32 | 24.4 | 13.3 | 102 | 17 | AO003312 |
| 33 | 24.4 | 13.3 | 125 | 13 | N72032 |
| 34 | 24.4 | 13.3 | 125 | 13 | B1003896 |
| 35 | 24.4 | 13.3 | 131 | 12 | BF361440 |
| 36 | 24.4 | 13.3 | 165 | 17 | AL753614 |
| 37 | 24.2 | 13.2 | 133 | 17 | AZ038645 |
| 38 | 24.2 | 13.2 | 143 | 9 | A1933911 |
| 39 | 24.2 | 13.2 | 146 | 13 | B1014386 |
| 40 | 24.2 | 13.2 | 148 | 13 | BG981038 |
| 41 | 24.2 | 13.2 | 153 | 9 | AA254114 |
| 42 | 24.2 | 13.2 | 167 | 9 | A1573113 |
| 43 | 24.2 | 13.2 | 168 | 12 | BF550916 |
| 44 | 24 | 13.0 | 110 | 13 | B1050851 |
| 45 | 24 | 13.0 | 159 | 9 | AJ394131 |

ALIGNMENTS

RESULT 1
AA645033 160 bp mRNA linear EST 28-OCT-1997
vs86f01.r1 Barstead mouse myotubes MRLB5 Mus musculus cDNA clone
IMAGE:1153177 5', mRNA sequence.

ACCESSION
AA645033
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Mammalia: Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 160)
Mairra, M., Hallier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE
JOURNAL
COMMENT

The WashU-HMMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:626385
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 132.
Location/Qualifiers

FEATURES

source

```
1..160
/organism="Mus musculus"
/strain="C3H"
/db_xref="taxon:10090"
/clone="IMAGE:115317"
/clone_1lb="Barstead mouse myotubes MPLRB5"
/cell_line="C2C12"
/lab_host="DH10B"
/Note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TCGTACGATCTGACGTGGAGGCGCCCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGATCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead. The C2C12 cell line
(available from ATCC, catalog # CRL-1772) differentiates
rapidly, forming contractile myotubes and producing
characteristic muscle proteins."
```

BASE COUNT

37 a 33 c 44 g 46 t

ORIGIN

Query Match 40.5%; Score 74.6; DB 9; Length 160;
Best Local Similarity 78.8%; Pred. No. 1.6e-13;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 15 GTACAGACAGTGTGTGTCCTGTGCTGCTTAACTTAAAGAGCGGTCTCAGA 74

Db 48 GTTCAGGTACAGACAGCGTTTCTGCTGCTTAACTTAAAGAGCGGTCTCAGA 107

QY 75 ATTGATAGACACGACAAACTGACGATGCTTGGATACCTCT 127

Db 108 ATTGATAGACACGACAAACTGCTGCTGCTTGGATGCTGCT 160

RESULT 2

AZ071494/c

LOCUS 155 bp DNA linear GSS 31-MAR-2000
DEFINITION RPCI-23-434112.TJ RPCI-23 Mus musculus genomic clone RPCI-23-434112
, DNA sequence.

ACCESSION AZ071494
VERSION AZ071494
KEYWORDS GI:7364479

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 155)
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akhmet,
B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,W., de Jong,P.
and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSSs: RPCI-23-434112.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
-9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

AUTHORS

TITLE
JOURNAL
COMMENT

Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 434 row: 1 column: 12
Seq primer: SP6

Class: BAC ends.
Location/Qualifiers
1..155
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-434112"
/clone_1lb="RPCI-23"
/sex="Female"
/lab_host="DH10B"

FEATURES

source

/Note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT

34 a 43 c 37 g 41 t

ORIGIN

Query Match 15.8%; Score 29; DB 17; Length 155;
Best Local Similarity 52.0%; Pred. No. 78;
Matches 65; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 59 AGCGGTCTTCAGATTGATTAAGACCATGACCAAACTGTGACGATTGGCTTGAG 118

Db 127 AGCGGTAGTCTCCACATATAAGACCTCCAGACAGCTGCAGCAATTTGTACCA 68

QY 119 TAACCTGTGTGCGTTCCTATTGACACAGAAATCAGCCTCATTCCTTAGTGAAG 178

Db 67 GCTTGGATGGATGTGCCAGTGGAGCTTAATAAGCCTCTCTCTAGTGCAGGAA 8

QY 179 CATTG 183

Db 7 CTTCG 3

RESULT 3

BE181187/c

LOCUS 153 bp mRNA linear EST 22-JUN-2000
DEFINITION CM2-HT0630-220300-125-h09 HT0630 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE181187
VERSION BE181187.1 GI:8660363

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 153)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.U.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

AUTHORS

TITLE
JOURNAL
MEDLINE
COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-CM2-HT0630-220>)
300-125-h09gt3-2000-03-22&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 9

Matches 43; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```

QY      68  TCTCAGGAATG  78
          |||||
Db      140 TCTCTGTATAG  150

```

| | | | |
|----|-----|--------------|-----|
| QY | 68 | TCTCAGAAITG | 78 |
| | | | |
| Db | 140 | TCTCTGTATTAG | 156 |

| RESULT 7 | | | | |
|------------|----------------|-----------|----------------------|-------------------------|
| LOCUS | R35649 | 142 bp | mRNA | linear |
| DEFINITION | SNMHDA020031SK | SmMw | Schistosoma | EST 07-SEP-1995 |
| ACCESSION | R35649 | | mansoni | CDNA 5', mRNA sequence. |
| VERSION | R35649.1 | GI:981309 | | |
| KEYWORDS | EST. | | Schistosoma mansoni. | |
| SOURCE | | | | |

| | | | | | |
|------------|---|-----------|------|--------|-----------------|
| LOCUS | R25649 | 142 bp | mRNA | linear | EST 07-SEP-1995 |
| DEFINITION | SNNHDA020031SK Smaw Schistosoma mansoni CDNA 5', mRNA sequence. | | | | |
| ACCESSION | R25649 | | | | |
| VERSION | R25649.1 | GI:981309 | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Schistosoma mansoni. | | | | |
| ORGANISM | Schistosoma mansoni | | | | |
| | Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; | | | | |
| | Strifeiidae; Schistosomatoidea; Schistosomatidae; Schistosoma. | | | | |

| REFERENCE | AUTHORS | TITLE |
|--------------------|----------------|---|
| 1 (bases 1 to 142) | Ridgers, I. L. | cDNAs from <i>Schistosoma mansoni</i> ; |

Journal
Unpublished (1995)
Contact: Ridgers, I.H.

Experimental Taxonomy Unit, Zoology
The Natural History Museum
Cromwell Road, South Kensington, London, SW7 5BD, U.K.
Tel: 01719389297

Schistosoma mansoni homology with emb|X71383|PCRPS5 P. carnea mRNA for 40S ribosomal protein S5
Seq primer: SK.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .142 |

```

/organism="Schistosoma mansoni"
/strain="Egyptian"
/db_xref="taxon:6183"
/clone_1lb="5maw"
/lab_host="Golden Hamster"
/notes="Vector: Bluescript; Site_1: EcoRI; Site_2: XhoI;
total mRNA was isolated from adult worms using the acid
guanidine thiocyanate method (Chomczynski and Sacchi, 1987
). Poly (A)+ RNA was purified by affinity chromatography
on oligo dT column (Aviv and Leder, 1972). The cDNA
expression library was constructed using the lambda Zap
cDNA cloning system (Stratagene-U01-Zap XR Gigapack II
cloning kit). The library was amplified (end titre=0.5 x
108 pfu/ml) and its efficiency tested for using known
S.mansoni probes (chorion and tropomyosin genes). The size
range of inserts for the library had 0.5 kb to 2.5 kb. The

```

| | | | | | |
|-----------------------|---|------------------|-----------|-------------|-----------|
| BASE COUNT | 43 a | 25 c | 28 g | 32 t | 14 others |
| ORIGIN | cDNA being directionally cloned between the EcoRI and the XhoI sites of the Bluescript phagemid. Bluescript phagemid was excised by mass in vivo excision of the library and inserted into E.coli XL1 Blue. The plasmid was purified using the Qiaagen QiaWell 8 system and sequenced with Tag dye-terminator chemistry (SK173 and T7 primers) for ABI 373A. The library was prepared by Hanan Abdel Hamed, Ain Shams Univ., Cairo and provided by Mohamed Saber , Theodore Billarz Research Institute, Cairo." | | | | |
| Query Match | 14.1%; | Score 26; | DB 14; | Length 142; | |
| Best Local Similarity | 51.8%; | Pred. No. 7e+02; | | | |
| Matches 44; | Conservative 0; | Mismatches 41; | Indels 0; | Gaps 0; | |

71 CAGATTATATAGACCACTGGCACAAAACGTGACGTTTGGATTAACCCTGTGTG 130
||||| ||||| ||| | |||||
20 CAGAAATCTTAGACACTTCAGAAAGNTTAAGCCGATNMANINCTNAITTCGGTTAANG 79 Db

QY 131 CGGTTCTATTGCACAGAAATCAGA 155
 ||| ||| | ||||| |
 Db 80 TAGTNCGAATTCCTATGCAATCATA 104

RESULT 8
BG257473/c 154 bp mRNA linear EST 13-FEB-2001
LOCUS
DEFINITION BG257473.1 NIH_MGC_92 Homo sapiens CDNA clone IMAGE:4508618 5', mRNA sequence.
ACCESSION BG257473
VERSION BG257473.1 GI:12767289
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 154)
NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@pds-riemail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
plate: L1AM10387 row: 1 column: 03
High quality sequence stop: 154.
Location/Qualifiers
1. 154
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4508618"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: PCMV-SPORE6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."

BASE COUNT 57 a 23 c 21 g 53 t

ORIGIN

Query Match 14.1%; Score 26; DB 12; Length 154;
Best Local Similarity 54.1%; Pred. No. 7.2e+02;
Matches 53; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 33 TGTTCCTGTCGCTACATTAAAGCGGCTTTCTCAGATTGATAGCCATGCA 92
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 152 TTTTCTTTTGTGATATATATCAATGTTATTTTAAAAAATATATAACATTCACA 93
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

QY 93 CAAACGTGTGACGATGGCTTGGAGTACCTGTGTG 130
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 92 TAACTGTGTACACATTTTGGGTGGAAAAAAGACATGTG 55
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 9
BH887597/c 142 bp DNA linear GSS 07-AUG-2002
LOCUS
DEFINITION BH887597.1 Leishmania major Friedlin BAC Library Leishmania major genomic clone LB01545a, DNA sequence.
ACCESSION BH887597
VERSION BH887597.1 GI:22132712
KEYWORDS GSS.
SOURCE Leishmania major.
ORGANISM Leishmania major
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
REFERENCE 1 (bases 1 to 142)
Myler, P.J., Vogt, C., Munden, H., Robertson, L., Sisk, E., Fazelinia, G., Aggarwal, G., Nelson, S., Seyler, A., Worthey, E., Stuart, K. and Ragland, M.

TITLE Leishmania major Friedlin BAC End Sequences
JOURNAL Unpublished (2002)
COMMENT Other-GSS: LB01545a.d_spc.1
Contact: Myler PJ
Seattle Biomedical Research Institute
4 Nickerson Street, Seattle, WA 98109-1651, USA
Tel: 206 284-8846
Fax: 206 284-0313
Email: mylerpj@sbri.org
Seq primer: T7
Class: BAC ends.

FEATURES
source
1. 142
/organism="Leishmania major"
/strain="Friedlin"
/db_xref="taxon:9664"
/clone="LB01545a"
/clone_lib="Leishmania major Friedlin BAC Library"
/lab_host="E. coli GeneHogs + TrifA"
/note="Vector: pCG270; Site 1: HindIII; Genomic DNA from Leishmania major Friedlin in agarose blocks was partially digested with HindIII, size selected, and ligated with HindIII-digested pCG270 vector DNA. 10368 clones were picked and arrayed in 384- and 96-well plates. Library construction and arraying was carried out by Resgen Corporation and clones and filters are available from them"

BASE COUNT 55 a 33 c 26 g 28 t

ORIGIN

Query Match 14.0%; Score 25.8; DB 17; Length 142;
Best Local Similarity 67.9%; Pred. No. 8.2e+02;
Matches 36; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 4 TCATCTTCAGTACAGACAGCTGTTGTGTCCTGTTGGCTAACCTTAA 56
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 66 TCAATCTTATGTTACAGGGGTGTGTGTGTATTTTGTCTATATGTAA 14
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 10
AV281230 178 bp mRNA linear EST 05-NOV-1999
LOCUS
DEFINITION AV281230 RIKEN full-length enriched, adult male testis (DH10B) Mus musculus CDNA clone 4933423009 3', mRNA sequence.
ACCESSION AV281230
VERSION AV281230.1 GI:6269267
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murine; Mus.
REFERENCE 1 (bases 1 to 178)
Kono, H., Aizawa, K., Akahira, S., Akiyama, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Harai, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, K., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomioka, N., Tsunoda, Y., Watanabe, S., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al. 1999)
Unpublished (1999)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>

This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
Seq primer: SP6
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..152
/organism="Homo sapiens"
/db_xref="GDB:7608182"
/db_xref="taxon:9606"
/clone="RPCT-11-28215"
/clone_lib="RPCT-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"

BASE COUNT 52 a 24 c 25 g 51 t

ORIGIN

Query Match 13.9%; Score 25.6; DB 17; Length 152;
Best Local Similarity 55.7%; Pred. No. 9.7e+02;
Matches 49; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 7 TTCTTCAGGTACAGACAGTGTGTGTCCTGTTGTGCTTAACCTTAAGAAGCGGTGT 66

Db 25 TTCTTTAAAAACGGAATGTAAATGTCTTCTCTATGAAGAAGATTAAGATTCACC 84

QY 67 TTCTCAGAAATTGATTAAGCATTCGCACA 94

Db 85 ATGATGATTATTATTAGACCCACAGACACA 112

RESULT 13

LOCUS BG980506 168 bp mRNA linear EST 12-JUN-2001
DEFINITION MR3-CN0147-300101-003-e03 CN0147 Homo sapiens CDNA, mRNA sequence.
ACCESSION BG980506
VERSION BG980506.1 GI:14383241
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 168)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagel,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

REFERENCE

AUTHORS

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/getttm2.pl?tl=MR3&cl=MR3-CN0147-
300101-003-e03&ts=2001-01-30&tl=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 167.
Location/Qualifiers

FEATURES

source

1..168
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CN0147"

/dev_stage="Adult"
/note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 40 a 42 c 30 g 56 t

ORIGIN

Query Match 13.8%; Score 25.4; DB 13; Length 168;
Best Local Similarity 54.9%; Pred. No. 1.2e+03;
Matches 50; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 64 TGTTCACAAATTGATTAACACATGCGACAAACTGTCGATTCCTTTGGAGTACC 123

Db 124 TATTGGTCACTGTGTGAGAGAGTTTTCACAAAAGTCAATGCGCTGTGGAAATGTG 65

QY 124 CTGTGTGCGTCTCTATTGCACAGAAATCAG 154

Db 64 AGGGAGTGGTGCGCTCCACACAGATTCAG 34

RESULT 14

LOCUS BBI66886 183 bp mRNA linear EST 29-JUN-2000
DEFINITION BBI66886 RIKEN full-length enriched, 16 days neonate thymus Mus
musculus cDNA clone A13090507 3' similar to AF057026 Rattus
norvegicus protein kinase KID-1 (Kid1) mRNA, mRNA sequence.
ACCESSION BBI66886
VERSION BBI66886.1 GI:8622827
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1 (bases 1 to 183)
Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koye,S., Kurihara,C., Kusakabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
,Y., Shigenoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,
,T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.,
Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
,M., Muramatsu,M. and Hayashizaki,Y.

REFERENCE

AUTHORS

TITLE

JOURNAL RIKEN Mouse ESTs (Kono,H., et al.)

COMMENT

Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermolabile and thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Katsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2003, 11:58:34 : Search time 999.224 Seconds
(without alignments)
1361.478 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 84

Sequence: 1 acttcattcttcaggtacag.....gttcctcagaattgataaga 84

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estha:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estov:*
- 5: em_estow:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_plo:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rpod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 84 | 100.0 | 439 | 10 | AM238830 xb34912.y |
| 2 | 84 | 100.0 | 464 | 12 | BF724951 bx10e07.y |
| 3 | 84 | 100.0 | 474 | 12 | BG698550 B02658576 |
| 4 | 84 | 100.0 | 491 | 12 | BF724196 bx02a08.y |
| 5 | 84 | 100.0 | 511 | 10 | AM247557 2819375.5 |
| 6 | 84 | 100.0 | 542 | 10 | BE264417 601191725 |

| | | | | | | | |
|----|------|-------|------|----|----------|----------|-----------|
| 7 | 84 | 100.0 | 715 | 12 | BG721839 | BG721839 | 602694439 |
| 8 | 84 | 100.0 | 740 | 13 | B116433 | B116433 | 602868855 |
| 9 | 84 | 100.0 | 746 | 13 | B1334702 | B1334702 | 602998683 |
| 10 | 84 | 100.0 | 768 | 12 | BG687640 | BG687640 | 602639251 |
| 11 | 84 | 100.0 | 775 | 12 | BG828196 | BG828196 | 602753612 |
| 12 | 84 | 100.0 | 776 | 13 | B1561125 | B1561125 | 603253647 |
| 13 | 84 | 100.0 | 803 | 13 | B1562083 | B1562083 | 603256666 |
| 14 | 84 | 100.0 | 806 | 13 | B1255914 | B1255914 | 602976536 |
| 15 | 84 | 100.0 | 809 | 12 | BG539839 | BG539839 | 602563723 |
| 16 | 84 | 100.0 | 813 | 12 | BE907823 | BE907823 | 601501935 |
| 17 | 84 | 100.0 | 865 | 14 | BE791785 | BE791785 | 601582154 |
| 18 | 84 | 100.0 | 879 | 14 | BQ216402 | BQ216402 | AGENCOURT |
| 19 | 84 | 100.0 | 882 | 14 | BQ218693 | BQ218693 | AGENCOURT |
| 20 | 84 | 100.0 | 891 | 9 | AL556377 | AL556377 | AL556377 |
| 21 | 84 | 100.0 | 915 | 14 | BQ431952 | BQ431952 | AGENCOURT |
| 22 | 84 | 100.0 | 947 | 9 | AL527764 | AL527764 | AL527764 |
| 23 | 84 | 100.0 | 975 | 14 | BQ055396 | BQ055396 | AGENCOURT |
| 24 | 84 | 100.0 | 1011 | 14 | BQ071045 | BQ071045 | AGENCOURT |
| 25 | 84 | 100.0 | 1026 | 13 | BM557410 | BM557410 | AGENCOURT |
| 26 | 84 | 100.0 | 1077 | 13 | B1251944 | B1251944 | 602952566 |
| 27 | 84 | 100.0 | 1121 | 14 | BQ067369 | BQ067369 | AGENCOURT |
| 28 | 84 | 100.0 | 1179 | 14 | BQ069475 | BQ069475 | AGENCOURT |
| 29 | 84 | 100.0 | 1310 | 14 | BQ898393 | BQ898393 | AGENCOURT |
| 30 | 84 | 100.0 | 1346 | 14 | BQ071343 | BQ071343 | AGENCOURT |
| 31 | 83 | 98.8 | 366 | 9 | AL561183 | AL561183 | AL561183 |
| 32 | 82.4 | 98.1 | 336 | 14 | BM745082 | BM745082 | K-EST0018 |
| 33 | 82.4 | 98.1 | 375 | 12 | BE766775 | BE766775 | 113-NT010 |
| 34 | 82.4 | 98.1 | 532 | 10 | AM161050 | AM161050 | au78c05.y |
| 35 | 82.4 | 98.1 | 540 | 10 | AW372346 | AW372346 | PM4-BP034 |
| 36 | 82.4 | 98.1 | 625 | 12 | BF529544 | BF529544 | 602044319 |
| 37 | 82.4 | 98.1 | 649 | 14 | BM832648 | BM832648 | K-EST0107 |
| 38 | 80.8 | 96.2 | 630 | 13 | B1826839 | B1826839 | 603075685 |
| 39 | 80.4 | 95.7 | 509 | 10 | BE265295 | BE265295 | 601193485 |
| 40 | 79.2 | 94.3 | 652 | 13 | B1764233 | B1764233 | 603045914 |
| 41 | 73 | 86.9 | 869 | 12 | BF571876 | BF571876 | 602076773 |
| 42 | 73 | 86.9 | 953 | 12 | BF528149 | BF528149 | 602042733 |
| 43 | 72.4 | 86.2 | 710 | 13 | B1559395 | B1559395 | 603253075 |
| 44 | 72 | 85.7 | 592 | 12 | BF978976 | BF978976 | 602147658 |
| 45 | 72 | 85.7 | 687 | 12 | BF91001 | BF91001 | 602251053 |

ALIGNMENTS

RESULT 1
LOCUS AM238830 439 bp mRNA linear EST 13-DEC-1999
DEFINITION xb34912.y1 NCI_CGAP_Lu31 Homo sapiens cDNA clone IMAGE:2578246 5',
mRNA sequence.
ACCESSION AM238830
VERSION AM238830.1 GI:6571296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 439)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail@nci.nih.gov
Tissue Procurement: ATCC cDNA Library Preparation: life
Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNW at:
www-bio.llnl.gov/bdrp/image/image.html
Seq primer: -408P from Glibco
High quality sequence stop: 425.
Location/Qualifiers

```

source
1. 439
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2578246"
/clone_lib="NCL_CGAP_Lu31"
/sex="male"
/dev_stage="fetal, 14 wk post-conception"
/lab_host="DH10B"
/note="Organ: lung, cell line: Vector: PCMV-SPORT6;
Site:1: EcoRV; Site:2: NotI; Cloned unidirectionally, no
5' adaptor. Primer: Oligo dt. Full-length library
constructed by Life Technologies."

BASE COUNT      120 a      85 c      107 g      127 t
ORIGIN
Query Match      100.0%; Score 84; DB 10; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAAGG 60
|||||
Db 35 ACTTCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAAGG 94
|||||

QY 61 CGGTGTTCTCAGATTGATAGA 84
|||||
Db 95 CGGTGTTCTCAGATTGATAGA 118
|||||

RESULT 2
LOCUS BF724951 464 bp mRNA linear EST 05-JAN-2001
DEFINITION bx10e07.y1 Human Iris cDNA (un-normalized, unamplified): BX Homo
sapiens cDNA clone bx10e07 5', mRNA sequence.
ACCESSION BF724951
VERSION BF724951.1 GI:12040870
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 464)
Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
NEIBANK: EST analysis and bioinformatics for ocular genomics
Invest. Ophthalmol. Vis. Sci. 41, (2000) In press
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 10 row: e column: 07
Seq primer: M13RPL reverse primer (ABI).
Location/Qualifiers
1. 464
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="bx10e07"
/clone_lib="Human Iris cDNA (un-normalized, unamplified):
BX"
/rissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMDH10B"
/note="Organ: Eye; Vector: PCMVSPORT6; Post-mortem iris
tissue was pooled from 10 individuals ranging in age from
4-80 years and RNA was extracted. From this pooled sample
an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A
directionally cloned cDNA library in the PCMVSPORT6 vector
was constructed at Life Technologies, essentially
following the protocols of the Superscript Plasmid System
full details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com). First
strand synthesis was carried out using a Not I

```

```

BASE COUNT      124 a      97 c      122 g      121 t
ORIGIN
Query Match      100.0%; Score 84; DB 12; Length 464;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAAGG 60
|||||
Db 58 ACTTCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAAGG 117
|||||

QY 61 CGGTGTTCTCAGATTGATAGA 84
|||||
Db 118 CGGTGTTCTCAGATTGATAGA 141
|||||

RESULT 3
LOCUS BG698550 474 bp mRNA linear EST 07-MAY-2001
DEFINITION 602658576F2 NCL_CGAP_Skn3 Homo sapiens cDNA IMAGE:4801602 5',
mRNA sequence.
ACCESSION BG698550
VERSION BG698550.1 GI:13965955
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 474)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps@emall.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM0684 row: h column: 19
High quality sequence start: 7
High quality sequence stop: 472.
Location/Qualifiers
1. 474
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4801602"
/clone_lib="NCL_CGAP_Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCL_CGAP Library."

BASE COUNT      123 a      94 c      122 g      135 t
ORIGIN
Query Match      100.0%; Score 84; DB 12; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAAGG 60
|||||
Db 80 ACTTCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTAAGG 139
|||||

QY 61 CGGTGTTCTCAGATTGATAGA 84
|||||

```

| | |
|---|---|
| ACCESSION | MRNA sequence. |
| VERSION | AM247557 |
| KEYWORDS | AM247557.1 GI:6590550 |
| SOURCE | EST. |
| ORGANISM | human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |
| REFERENCE | 1 (bases 1 to 511) |
| AUTHORS | NIH-MGC http://mgc.ncl.nih.gov/ . |
| TITLE | National Institutes of Health, Mammalian Gene Collection (MGC) |
| JOURNAL | Unpublished (1999) |
| COMMENT | Other_ESTRs: 2819375.3prime Contact: Robert Strausberg, Ph.D. Email: cgaaps+email.nih.gov Tissue Procurement: DCTD/DTP CDNA Library Preparation: Ling Hong/Rubin laboratory CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Berkeley MGC sequencing project clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: www.bld.lnl.nih.gov/db/rp/image/image.html Base Calling / Quality Scores: PHRD from University of Washington Genome Center. Vector Trimming: cross-match from University of Washington Genome Center PHRAP suite. Poly-T identification: patmatch.pl from Berkeley Drosophila genome project. University of Washington Genome Center: http://www.genome.washington.edu Plate: LILCM1 row: G column: 24 High quality sequence stop: 464. location/Qualifiers 1..511 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image="IMAGE:2819375" /clone_id="NIH_MGC_7" /tissue_type="small cell carcinoma" /cell_line="MGC3" /lab_host="DH10B (phage-resistant)" /note="Organ: lung; Vector: pOTR7; Site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." BASE COUNT 135 a 109 c 141 g 126 t ORIGIN |
| Query Match | 100.0%; Score 84; DB 10; Length 511; |
| Best Local Similarity | 100.0%; P-adj. No. 1.1e-16; |
| Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| QY 1 ACTTCATCTCTCAGGTACAGACAGTGTTGCGTCCCTGGTGGCTAACATTAAAGAAG 60 | |
| Ddb 71 ACTTCATCTCTCAGGTACAGACAGTGTTGCGTCCCTGGTGGCTAACATTAAAGAAG 130 | |
| QY 61 CGGTGTTCTCAGAAATGATAAGA 84 | |
| Ddb 131 CGGTGTTCTCAGAAATGATAAGA 154 | |
| RESULT 6 BE264417 | BE264417 |
| LOCUS DEFINITION | 601191725F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3535398 5', mRNA linear EST 13-JUL-2000 |
| ACCESSION VERSION | BE264417 BE264417.1 GI:9137973 |
| KEYWORDS | mRNA sequence. |
| SOURCE | EST. |
| ORGANISM | human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |

REFERENCE 1 (bases 1 to 542)
 AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Plate: L1CM216 row: b column: 07
 High quality sequence stop: 542.
 Location/Qualifiers

FEATURES
 source 1..542
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:353598"
 /clone_lib="NIH_MGC_7"
 /issue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 154 a 121 c 142 g 125 t
 ORIGIN

Query Match 100.0%; Score 84; DB 10; Length 542;
 Best Local Similarity 100.0%; Pred. No. 1.1e-16;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTCCTGTTGGCTAACTTAAGAAG 60
 |||||||
 Db 37 ACTTCATTTCTTCAGTACAGACAGTGTGTCCTGTTGGCTAACTTAAGAAG 96

QY 61 CGGTGTTCTCAGAAATTGATAGA 84
 |||||||
 Db 97 CGGTGTTCTCAGAAATTGATAGA 120

RESULT 7
 BG721839 715 bp mRNA linear EST 08-MAY-2001
 LOCUS 602694439F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4826697 5',
 DEFINITION mRNA sequence.
 ACCESSION BG721839
 VERSION BG721839.1 GI:14001026
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 715)
 AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshitaki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: L1AM10741 row: n column: 10
 High quality sequence stop: 708.
 Location/Qualifiers

FEATURES
 source 1..715
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/clone="IMAGE:4826697"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (ggcgaag); Oligo-dT primed using primer 5'-TTTTTTTTTTT-3', size-selected for average insert size 2.2 kb and normalized to 50x. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI/NHRI, National Institutes of Health). Note: this is a NIH_MGC library."
 BASE COUNT 199 a 154 c 189 g 173 t
 ORIGIN

Query Match 100.0%; Score 84; DB 12; Length 715;
 Best Local Similarity 100.0%; Pred. No. 1.1e-16;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTCCTGTTGGCTAACTTAAGAAG 60
 |||||||
 Db 80 ACTTCATTTCTTCAGTACAGACAGTGTGTCCTGTTGGCTAACTTAAGAAG 139

QY 61 CGGTGTTCTCAGAAATTGATAGA 84
 |||||||
 Db 140 CGGTGTTCTCAGAAATTGATAGA 163

RESULT 8
 B1116433 740 bp mRNA linear EST 26-JUN-2001
 LOCUS 60268855F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5017525 5',
 DEFINITION mRNA sequence.
 ACCESSION B1116433
 VERSION B1116433.1 GI:14567334
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 740)
 AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCDP/DRP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
 Plate: L1CM1829 row: n column: 14
 High quality sequence stop: 698.
 Location/Qualifiers

FEATURES
 source 1..740
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5017525"
 /clone_lib="NIH_MGC_7"
 /issue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 206 a 162 c 199 g 173 t
 ORIGIN

Query Match 100.0%; Score 84; DB 13; Length 740;
Best Local Similarity 100.0%; Pred. No. 1,1e-16;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATCTTCAGGTACAGACAGTGTTCGTGCTTCCTGTTGGCTAACTTTAAGAAG 60
|||||
Db 25 ACTTCATCTTCAGGTACAGACAGTGTTCGTGCTTCCTGTTGGCTAACTTTAAGAAG 84
|||||

QY 61 CGGTGTTTCTCAGAAATGATAGA 84
|||||
Db 85 CGGTGTTTCTCAGAAATGATAGA 108
|||||

RESULT 9
BI334702 746 bp mRNA linear EST 30-JUL-2001
LOCUS 602998683F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5140628 5',
DEFINITION mRNA sequence.
ACCESSION BI334702
VERSION BI334702.1 GI:15019359
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 746)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM11346 row: f column: 21
High quality sequence stop: 744.
Location/Qualifiers
1..746
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5140628"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 210 a 158 c 202 g 176 t

ORIGIN
100.0%; Score 84; DB 13; Length 746;
Best Local Similarity 100.0%; Pred. No. 1,1e-16;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATCTTCAGGTACAGACAGTGTTCGTGCTTCCTGTTGGCTAACTTTAAGAAG 60
|||||
Db 35 ACTTCATCTTCAGGTACAGACAGTGTTCGTGCTTCCTGTTGGCTAACTTTAAGAAG 94
|||||

QY 61 CGGTGTTTCTCAGAAATGATAGA 84
|||||
Db 95 CGGTGTTTCTCAGAAATGATAGA 118
|||||

RESULT 10
BG687640 768 bp mRNA linear EST 01-MAY-2001
LOCUS 602639251F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4762356 5',
DEFINITION

mRNA sequence.
BG687640
VERSION BG687640.1 GI:13919037
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 768)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMI616 row: e column: 13
High quality sequence stop: 716.
Location/Qualifiers
1..768
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4762356"
/clone_lib="NIH_MGC_59"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgcctggcc); Site_2: SfiI (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCGCATATGAGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGAGCGGCGCATG-RT(30)BN-3'
(where B = A, C or G and N = A, C, G or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH-MGC
Library."

BASE COUNT 214 a 163 c 212 g 179 t

ORIGIN
100.0%; Score 84; DB 12; Length 768;
Best Local Similarity 100.0%; Pred. No. 1,1e-16;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATCTTCAGGTACAGACAGTGTTCGTGCTTCCTGTTGGCTAACTTTAAGAAG 60
|||||
Db 58 ACTTCATCTTCAGGTACAGACAGTGTTCGTGCTTCCTGTTGGCTAACTTTAAGAAG 117
|||||

QY 61 CGGTGTTTCTCAGAAATGATAGA 84
|||||
Db 118 CGGTGTTTCTCAGAAATGATAGA 141
|||||

RESULT 11
BG828196 775 bp mRNA linear EST 22-MAY-2001
LOCUS 602753612F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4906394 5',
DEFINITION mRNA sequence.
ACCESSION BG828196
VERSION BG828196.1 GI:14175796
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 775)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLM1809 row: 9 column: 03
 High quality sequence stop: 758.

FEATURES

Location/Qualifiers

1..775

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1b="NIH_MGC_97"
 /issue_type="rhabdomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
 Site_2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 215 a 167 c 213 g 180 t
 ORIGIN

Query Match 100.0%; Score 84; DB 12; Length 775;
 Best Local Similarity 100.0%; Pred. No. 1..1e-16;

Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTCTGTGCTTGGCTTAATTGAAG 60
 |||||||
 DB 61 ACTTCATTTCTTCAGTACAGACAGTGTGTTCTGTGCTTGGCTTAATTGAAG 120
 |||||||
 QY 61 CGGTGTTTCGAGATTGATTAAGA 84
 |||||||
 DB 121 CGGTGTTTCGAGATTGATTAAGA 144
 |||||||

RESULT 12 776 bp mRNA linear EST 05-SEP-2001
 BI561125
 LOCUS 603253647F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296306 5',
 DEFINITION mRNA sequence.
 BI561125
 ACCESSION BI561125.1 GI:15448439
 VERSION
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE NIH-MGC <http://mgc.ncl.nih.gov/>
 1 (bases 1 to 776)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM11749 row: m column: 11
 High quality sequence stop: 774.

FEATURES

source

Location/Qualifiers

1..776

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1b="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptPR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag
) Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.2 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI/NIH), National
 Institutes of Health). Note: this is a NIH-MGC library."
 BASE COUNT 213 a 168 c 216 g 179 t
 ORIGIN

Query Match 100.0%; Score 84; DB 13; Length 776;
 Best Local Similarity 100.0%; Pred. No. 1..1e-16;

Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTTCATTTCTTCAGTACAGACAGTGTGTTCTGTGCTTGGCTTAATTGAAG 60
 |||||||
 DB 65 ACTTCATTTCTTCAGTACAGACAGTGTGTTCTGTGCTTGGCTTAATTGAAG 124
 |||||||
 QY 61 CGGTGTTTCGAGATTGATTAAGA 84
 |||||||
 DB 125 CGGTGTTTCGAGATTGATTAAGA 148
 |||||||

RESULT 13 803 bp mRNA linear EST 05-SEP-2001
 BI562083
 LOCUS 60325666F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298785 5',
 DEFINITION mRNA sequence.
 BI562083
 ACCESSION BI562083.1 GI:15449409
 VERSION
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE NIH-MGC <http://mgc.ncl.nih.gov/>
 1 (bases 1 to 803)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM11756 row: d column: 18
 High quality sequence stop: 766.

FEATURES

source

1..803

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1b="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptPR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag
) Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.2 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NIHRI, National Institutes of Health). Note: this is a NIH-MGC library.

| | | | | |
|------------|-------|-------|-------|-------|
| BASE COUNT | 217 a | 181 c | 220 g | 185 t |
| ORIGIN | | | | |

| | | | | |
|-----------------------|---------|--------------|----------|---------------|
| Query Match | 100.0% | Score 84: | DB 13: | Length 803; |
| Best Local Similarity | 100.0%: | Pred. No. | 1.1e-16; | |
| Matches | 84; | Conservative | 0; | Mismatches 0; |
| | | | | Indels 0; |
| | | | | Gaps 0; |

```
Oy      1 ACCTCATTTCACGAGTACAGACAGTGTTGTGGTCCCGTTGIGACTAATTAAGAAG   60  
        |||  
Db     83 ACTTCATTTCACGAGTACAGACAGTGTTGTGCCCTCAACTTAAGAAG    142
```

| | | | | |
|------------|-------|-------|-------|-------|
| BASE COUNT | 222 a | 178 c | 222 g | 184 t |
| ORIGIN | | | | |

| | | | | |
|-----------------------|----------------|--------------|----------|------------|
| Query Match | 100.0% | Score 84 | DB 13 | Length 806 |
| Best Local Similarity | 100.0% | Pred. No. | 1.1e-16 | |
| Matches 84 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |

QY 1 ACTTCATTCTTCAGGTACAGACAGTGTGTTGTTCCGTGGCTAACTTTAAGAAG 60
DB 60 ACTTCATTCTTCAGGTACAGACAGTGTGTTGTTCCGTGGCTAACTTTAAGAAG 119

| | | | | |
|------------|--|--------|------|-----------------|
| RESULT 15 | | | | |
| BG539839 | BG539839 | 809 bp | mRNA | linear |
| LOCUS | | | | |
| DEFINITION | 602565723nt NIH.MGC.77 Homo sapiens CDNA clone IMAGE:4688452 5', | | | EST 03-APR-2001 |

Db 106 CGGTGTTCTCAGAA TTGATAGA 129

Search completed: February 16, 2003, 01:20:11
job time : 1003.22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: February 16, 2003, 02:51:59 ; Search time 24.7612 Seconds

(without alignments)
1397.991 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 152

Sequence: 1 actcattcttcagtgtagcag.....gtttcagcagattgtaaga 84

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame+np.model -runat_12022003.170353.8526/app_query.fasta.1.590
-O=/cgn2_1/USPTO-spool/US0939293/runat_12022003.170353.8526/app_query.fasta.1.590
-DB=SPTRMBL_21 -QFMT=fastcan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45
-DOCAALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0939293@cgn.1.1.51.0runat_12022003.170353.8526 -NCP=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTRMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Query Length | DB ID | Description |
|------------|-------------|--------------|-------|-------------------------------|
| 1 | 87 | 57.2 | 157 | 11 Q8R1D8 Q8R1D8 mus musculus |

| | | | | | | |
|----|------|------|------|----|--------|--------------------|
| 2 | 55 | 36.2 | 445 | 5 | Q9NF64 | Q9NF64 caenorhabd |
| 3 | 55 | 36.2 | 449 | 10 | Q9S9V9 | Q9S9V9 arabidopsis |
| 4 | 54.5 | 35.9 | 687 | 10 | Q49728 | Q49728 arabidopsis |
| 5 | 52 | 34.2 | 160 | 12 | Q91FW7 | Q91FW7 chilo iride |
| 6 | 51.5 | 33.9 | 564 | 6 | Q9TTS4 | Q9TTS4 bos taurus |
| 7 | 51.5 | 33.9 | 1099 | 5 | Q9VMS5 | Q9VMS5 drosophila |
| 8 | 51.5 | 33.9 | 5146 | 6 | Q8SPM4 | Q8SPM4 bos taurus |
| 9 | 51 | 34.9 | 153 | 5 | P91741 | P91741 hydra atten |
| 10 | 51 | 34.9 | 268 | 16 | Q9PC68 | Q9PC68 xyrella fas |
| 11 | 51 | 34.9 | 675 | 8 | Q32213 | Q32213 eriomastax |
| 12 | 50.5 | 34.6 | 4834 | 4 | Q95714 | Q95714 mus musculi |
| 13 | 50.5 | 34.6 | 4836 | 11 | Q88473 | Q88473 mus musculi |
| 14 | 50 | 32.9 | 241 | 10 | Q9FHB4 | Q9FHB4 arabidopsis |
| 15 | 50 | 34.2 | 299 | 5 | Q9VU37 | Q9VU37 drosophila |
| 16 | 50 | 34.2 | 301 | 8 | Q9GCN2 | Q9GCN2 pictis pauc |
| 17 | 50 | 32.9 | 676 | 2 | Q8VTT3 | Q8VTT3 pseudomonas |
| 18 | 50 | 32.9 | 741 | 4 | Q9NZY6 | Q9NZY6 homo sapien |
| 19 | 50 | 34.2 | 744 | 8 | Q32671 | Q32671 nassauvia g |
| 20 | 50 | 34.2 | 806 | 5 | P91808 | P91808 strongyloce |
| 21 | 49.5 | 33.9 | 621 | 10 | Q9LYP5 | Q9LYP5 arabidopsis |
| 22 | 49.5 | 33.9 | 763 | 16 | Q929L6 | Q929L6 listeria in |
| 23 | 49.5 | 33.9 | 763 | 16 | Q8Y5B2 | Q8Y5B2 listeria mo |
| 24 | 49.5 | 33.9 | 772 | 6 | Q95JG9 | Q95JG9 sus scrofa |
| 25 | 49 | 32.2 | 274 | 6 | Q95L51 | Q95L51 capra hircu |
| 26 | 49 | 32.2 | 289 | 16 | Q9BET4 | Q9BET4 rhizobium l |
| 27 | 49 | 32.2 | 387 | 4 | Q8TCW4 | Q8TCW4 homo sapien |
| 28 | 49 | 32.2 | 418 | 13 | Q8UWM4 | Q8UWM4 brachydanto |
| 29 | 49 | 33.6 | 495 | 5 | Q9VJB3 | Q9VJB3 drosophila |
| 30 | 49 | 32.2 | 590 | 4 | Q96G63 | Q96G63 homo sapien |
| 31 | 49 | 32.2 | 632 | 10 | Q04537 | Q04537 arabidopsis |
| 32 | 49 | 32.2 | 636 | 4 | Q9H9B6 | Q9H9B6 homo sapien |
| 33 | 49 | 32.2 | 637 | 5 | Q45915 | Q45915 caenorhabd |
| 34 | 49 | 32.2 | 695 | 4 | Q9H062 | Q9H062 homo sapien |
| 35 | 49 | 33.6 | 744 | 8 | Q31817 | Q31817 peretia mic |
| 36 | 49 | 32.2 | 747 | 4 | Q96AE5 | Q96AE5 homo sapien |
| 37 | 49 | 32.2 | 747 | 4 | Q9NVG7 | Q9NVG7 homo sapien |
| 38 | 49 | 33.6 | 1136 | 5 | Q8MSV8 | Q8MSV8 plasmodium |
| 39 | 48.5 | 33.2 | 172 | 2 | Q9P532 | Q9P532 escherichia |
| 40 | 48.5 | 31.9 | 209 | 4 | Q96MW8 | Q96MW8 homo sapien |
| 41 | 48.5 | 33.2 | 213 | 2 | Q9RC24 | Q9RC24 bacillus sp |
| 42 | 48.5 | 33.2 | 213 | 16 | Q8YNC4 | Q8YNC4 anabaena sp |
| 43 | 48.5 | 33.2 | 401 | 16 | Q8UHHO | Q8UHHO agrobacteri |
| 44 | 48.5 | 33.2 | 756 | 4 | Q9BWK0 | Q9BWK0 homo sapien |
| 45 | 48 | 32.9 | 210 | 10 | Q8Z350 | Q8Z350 arabidopsis |

ALIGNMENTS

RESULT 1

| ID | Q8R1D8 | PRELIMINARY; | PRT; | 157 AA. |
|----|--|--------------|-----------|-------------------------|
| AC | Q8R1D8; | | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Created) | | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last sequence update) | | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) | | | |
| DE | Similar to RIKEN CDNA 0610041G12 gene. | | | |
| OS | Mus musculus (Mouse). | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OX | NCBI_TaxID=10090; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=EYE; | | | |
| RA | Strusberg R.; | | | |
| RL | Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; BC024780; AAH24780.1; - | | | |
| SO | SEQUENCE | 157 AA; | 17799 MW; | 0F67319F05EAC6E7 CRC64; |

Alignment Scores: 1.9e-05 Length: 157
Pred. No.: 87.00 Matches: 19
Score: 77.78% Conservative: 2
Percent Similarity: 70.37% Mismatches: 4

```

Query Match: 57.24% Indels: 2
DB: 11 Gaps: 1
US-09-939-293-1_COPY_56_139 (1-84) x Q9NF64 (1-157)
OY 4 TCATCTTCAGGTACAGACAGTGTGTGTCCTGTCGCTTACCTTAAGACGCG 63
DB 14 SerLeuPheArgTyrArgGlnArgPhe-----ProValLeuAlaSerLysArg 31
OY 64 TGTTCCTCAGAAATGATAGA 84
DB 32 CysPheSerGlnLeuLeuLys 38

RESULT 2
Q9NF64 PRELIMINARY; PRT; 445 AA.
ID Q9NF64;
AC Q9NF64;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Y105C5A.24 protein.
GN Y105C5A.24.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Feloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RA SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; AL117193; CAB55004.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF000069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 445 AA; 51034 MW; 5D461D6140D18F CRC64;

Alignment Scores:
Pred. No.: 3.5 Length: 445
Score: 55.00 Matches: 10
Percent Similarity: 64.00% Conservative: 6
Best Local Similarity: 40.00% Mismatches: 9
Query Match: 36.18% Indels: 0
Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x Q9NF64 (1-445)
OY 1 ACTTCATCTTCAGGTACAGACAGTGTGTGTCCTGTCGCTTACCTTAAGACG 60
DB 223 ThrArgAsnLeuArgProHisGlnLeuGluCysAsnProIleLeuSerAsnPhetyrLys 242
OY 61 CGGTGTTCTCAGAA 75
DB 243 ArgCysTrpSerAsp 247

RESULT 3
Q9S9V9 PRELIMINARY; PRT; 449 AA.
ID Q9S9V9;
AC Q9S9V9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE T1J24.2 protein (AT4G05500 protein).

```

```

GN T1J24.2 OR AT4G05500.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN-CV. COLUMBIA;
RA Ali J., Bauer C., Nguyen C., Duckels G.;
RT "The sequence of A. thaliana T1J24."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA WashU;
RT "The A. thaliana Genome Sequencing Project";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Waterston R.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF147263; AAD48965.1; -.
DR EMBL; AL161503; CAB81092.1; -.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00236; FBOX; 1.
DR PROSITE; PS50181; FBOX; 1.
SQ SEQUENCE 449 AA; 51108 MW; 8EFA04E4347718B6 CRC64;

Alignment Scores:
Pred. No.: 3.5 Length: 449
Score: 55.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 36.18% Indels: 0
Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x Q9S9V9 (1-449)
OY 19 AGACAGTGTGTGTCCTGTCGCTTACCTTAAGACGCGTGTCTCAGAAATG 78
DB 393 ArgGlnCysPheAsnIleAsnLeuValGlyAspLeuLysLysArgCysPheGluArgIle 412

RESULT 4
O49728 PRELIMINARY; PRT; 687 AA.
ID O49728;
AC O49728;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Receptor serine/threonine kinase-like protein.
GN T9A21.100 OR AT4G18250.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RA SEQUENCE FROM N.A.
RA Beyer M., Murphy G., Drost L., Hall C., Hudson S., Ridley P.,
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

```

[2]
 RP SEQUENCE FROM N.A.
 RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (Mar-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (Mar-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AL021713; CAB16797.1; -;
 DR EMBL: AL161548; CAB78827.1; -;
 DR HSSP: P02883; 1THW.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR004040; STT_pkinase.
 DR InterPro: IPR001938; Thaumatin.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00314; thaumatin; 2.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR ProDom: PD001321; Thaumatin; 1.
 DR SMART: SM00221; STYKc; 1.
 DR SMART: SM00205; THN; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00316; THAUMATIN; UNKNOWN; 1.
 KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 687 AA; 77096 MW; 3DE557112AF2C84 CRC64;

Alignment Scores:

| Pred. No.: | 4.25 | Length: | 687 |
|------------------------|--------|---------------|-----|
| Score: | 54.50 | Matches: | 5 |
| Percent Similarity: | 65.22% | Conservative: | 10 |
| Best Local Similarity: | 43.48% | Mismatches: | 7 |
| Query Match: | 35.86% | Indels: | 1 |
| DB: | 10 | Gaps: | 1 |

US-09-939-293-1_COPY_56_139 (1-84) x 049728 (1-687)

QY 1 ACTCATCTCTCAGTACAGACAG--TCTTGTGTCTTCCTGTGCTTAACTTAA 57
 |||:::||||:||||:||||| ||| ||| |||||
 DB 240 ThnAsnTryPheLysTyrLysGlnThrCysLysProThrProtyrValGlnAsnPhenap 259
 QY 58 AAGCGGTGT 66
 :::: |||
 DB 260 ArgAlaCys 262

RESULT 5
 091FW7 PRELIMINARY; PRT; 160 AA.
 ID 091FW7;
 AC 091FW7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 203L.
 OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).
 OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
 OX NCBI_TaxID=10488;
 RN NCBL_TaxID=10488;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Dellus H., Darai G., Fluegel R.M.;
 RT "DNA analysis of insect iridescent virus 6: evidence for circular
 RT permutation and terminal redundancy.";
 RL J. Virol. 49:609-614(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86174607; PubMed=395991;
 RA Lorchacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;
 RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis
 RT in mice.";
 RL Med. Microbiol. Immunol. 175:43-53(1986).
 RN [3]
 RP SEQUENCE FROM N.A.

RX MEDLINE=87321126; PubMed=2820141;
 RA Schiltzler P., Soltau J.B., Fischer M., Reiser H., Scholz J.,
 RA Dellus H., Darai G.;
 RT "Molecular cloning and physical mapping of the genome of insect
 RT iridescent virus type 6: further evidence for circular permutation of
 RT the viral genome.";
 RL Virology 160:66-74(1987).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89073752; PubMed=3201750;
 RA Fischer M., Schiltzler P., Dellus H., Darai G.;
 RT "Identification and characterization of the repetitive DNA element in
 RT the genome of insect iridescent virus type 6.";
 RL Virology 167:485-496(1988).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92196996; PubMed=1549908;
 RA Handermann M., Schiltzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,
 RA Darai G.;
 RT "Identification and mapping of origins of DNA replication within the
 RT DNA sequences of the genome of insect iridescent virus type 6.";
 RL Virus Genes 6:19-32(1992).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93118242; PubMed=1475907;
 RA Sonntag K.C., Darai G.;
 RT "Characterization of the third origin of DNA replication of the genome
 RT of insect iridescent virus type 6.";
 RL Virus Genes 6:333-342(1992).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93260401; PubMed=8492091;
 RA Stowasser R., Raab K., Schiltzler P., Janssen W., Darai G.;
 RT "Identification of the gene encoding the major capsid protein of
 RT insect iridescent virus type 6 by polymerase chain reaction.";
 RL J. Gen. Virol. 74:873-879(1993).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94167241; PubMed=8121799;
 RA Schiltzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
 RA Dellus H., Darai G.;
 RT "Identification of genes encoding zinc finger proteins, non-histone
 RT chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
 RT in the genome of Chilo iridescent virus.";
 RL Nucleic Acids Res. 22:158-166(1994).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94353641; PubMed=8073636;
 RA Sonntag K.C., Schiltzler P., Koonin E.V., Darai G.;
 RT "Chilo iridescent virus encodes a putative helicase belonging to a
 RT distinct family within the 'DEAD/H' superfamily: implications for the
 RT evolution of large DNA viruses.";
 RL Virus Genes 8:151-158(1994).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95213160; PubMed=7698884;
 RA Sonntag K.C., Schiltzler P., Janssen W., Darai G.;
 RT "Identification of the primary structure and the coding capacity of
 RT the genome of insect iridescent virus type 6 between the genome
 RT coordinates 0.310 and 0.347 (7990 bp).";
 RL Intervirology 37:287-297(1994).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94292906; PubMed=8021587;
 RA Schiltzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
 RA Koonin E.V., Darai G.;
 RT "Insect iridescent virus type 6 encodes a polypeptide related to the
 RT largest subunit of eukaryotic RNA polymerase II.";
 RL J. Gen. Virol. 75:1557-1567(1994).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98141693; PubMed=9482589;
 RA Bahr U., Tidona C.A., Darai G.;

| | |
|----|--|
| DR | SMART: SH00216; WMD: 1. |
| FT | NON_TER 1 |
| FT | NON_TER 564 |
| SO | SEQUENCE 564 AA; 60338 MW; 2E22D4242F9BBE7C CRC64; |

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-----|
| Pred. No.: | 13.2 | Length: | 564 |
| Score: | 51.50 | Matches: | 10 |
| Percent Similarity: | 68.42% | Conservative: | 3 |
| Best Local Similarity: | 52.63% | Mismatches: | 5 |
| Query Match: | 33.88% | Indels: | 1 |
| DB: | 6 | Gaps: | 1 |

US-09-939-293-1_COPY_56_139 (1-84) x O9YTS4 (1-564)

Oy 19 AGACAGTGTGTTGGTCCTGGTGCGTAACCTATGAAGACGGTGTCGCAGA 75
||| ||||||| ||||||:|||| :||: ||||||| |||
Db 443 Argaspsycystecys---ProvalLeuAlaLaLAtyrLarlgargcysAlaIngInlu 460

RESULT 7

| | | | |
|----|--|--------------|---------------|
| ID | O9VMS5 | PRELIMINARY: | PRT; 1099 AA. |
| AC | O9VMS5: | | |
| DT | 01-MAY-2000 (TREMBLrel. 13, Created) | | |
| DT | 01-May-2000 (TREMBLrel. 13, Last sequence update) | | |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last annotation update) | | |
| DE | CG14030 protein (LPD2858p). | | |
| GN | CG14030. | | |
| OS | Drosophila melanogaster (Fruit fly). | | |
| OC | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; | | |
| OC | Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; | | |
| OX | Ephydroidea; Drosophilidae; Drosophila. | | |
| RN | NCBI_TaxID=7227; | | |
| RP | [1] | | |
| RC | SEQUENCE FROM N.A. | | |
| RX | STRAIN=BKERELET; | | |
| RX | MEDLINE=20196006; PubMed=10731132; | | |
| RA | Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D., | | |
| RA | Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., | | |
| RA | George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., | | |
| RA | Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., | | |
| RA | Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D., | | |
| RA | Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G., | | |
| RA | Abbil J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D., | | |
| RA | Ballew R.M., Basu A., Bakendole J., Bayraktaroglu L., Beasley E.M., | | |
| RA | Beecon K.Y., Benos P.V., Bernan B.P., Bhattacharya D., Bolshakov S., | | |
| RA | Botkova D., Botchan M.R., Bouck J., Broksstein P., Brotlier P., | | |
| RA | Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., | | |
| RA | Cherry J.M., Cawley S., Dahlke G., Davenport L.B., Davies P., | | |
| RA | De Pallos B., Delcher A., Deng Z., Mayas A.D., Dew I., Dietz S.M., | | |
| RA | Dodson K., Doug L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., | | |
| RA | Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W. | | |
| RA | Foster C., Gabriellel A.E., Gary N.S., Gelbart W.M., Glasser K., | | |
| RA | Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., | | |
| RA | Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J., | | |
| RA | Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., | | |
| RA | Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., | | |
| RA | Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., | | |
| RA | Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., | | |
| RA | Lin X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., | | |
| RA | Merkulov G., Mishina N.V., Mobarry C., Morris J., Mosherel A., | | |
| RA | Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., | | |
| RA | Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacle J.M., | | |
| RA | Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G., | | |
| RA | Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., | | |
| RA | Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., | | |
| RA | Spyker E., Spradling A.C., Stapleton M., Strong R., Sun E., | | |
| RA | Svirskas R., Tector C., Turner R., Venler E., Wang A.H., Wang X., | | |
| RA | Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J., | | |
| RA | Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., | | |
| RA | Ye Y., Yen R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L., | | |
| RA | Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., | | |
| RA | Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; | | |


```

RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franco S.G., Franco M.C., Frohme M., Furlan L.R.,
RA Ganier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madela A.M.B.N., Madela H.M.F., Matsukuma A.Y.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E.Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A.Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Valada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AE004011; AAF84719.1; -.
DR InterPro: IPR001130; TatD.DNase.
DR Pfam: PF01026; TatD.DNase; 1.
DR PROSITE; PS01091; TATD_3; UNKNOWN_1.
RW Complete proteome.
SQ SEQUENCE 268 AA; 30412 MW; 4B1D134B81CD21B5 CRC64;

Alignment Scores:
Pred. No.: 15.9 Length: 268
Score: 51.00 Matches: 9
Percent Similarity: 61.90% Conservative: 4
Best Local Similarity: 42.86% Mismatches: 8
Query Match: 34.93% Indels: 0
DB: Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x Q9PC68 (1-268)
QY 83 CTATCAATTCGTGAGAAACCGCTTCTTAAGTTAGCCACAGAACACACACAC 24
Db 48 LeuLeuLeuAlaGlnArgHisProClyLeuLeuTyralaThrAlaGlyValHisProHis 67
QY 23 TGT 21
Db 68 Cys 68

RESULT 11
032213 PRELIMINARY: PRT; 675 AA.
AC Q32213.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NADH dehydrogenase subunit (Fragment).
GN NDHF.
OS Eremomastax spectiosa.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eumasterids I; Lamiales; Acanthaceae; Eremomastax.
OX NCBI_TaxId=37794;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Scotland R.W., Sweeney J.A., Reeves P.A., Oimstead R.G.;
RT "Higher level systematics of Acanthaceae determined by chloroplast DNA
RT sequences.";
RL Am. J. Bot. 0:0-0(1995).
CC -1- CATALYTIC ACTIVITY: NADH + PLASTOQUINONE = NAD(+) + PLASTOQUINOL.
DR EMBL; U12659; AAA61723.1; -.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam; PF00361; oxidored_q1; 1.
DR Pfam; PF01010; oxidored_q1_C; 1.
DR Pfam; PF00662; oxidored_q1_N; 1.
KW Chloroplast; NAD; Oxidoreductase; Plastoquinone.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 675 AA; 76873 MW; 943D9622D98FA693 CRC64;

Alignment Scores:
Pred. No.: 16 Length: 675
Score: 51.00 Matches: 9
Percent Similarity: 68.42% Conservative: 4
Best Local Similarity: 47.37% Mismatches: 6
Query Match: 34.93% Indels: 0
DB: Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x Q32213 (1-675)
QY 63 CCGCTTCTTAAGTTAGCCACAGAACACACACACACACACACACACACAC 7
Db 490 ProPhellelSileuTySerArgAsnLysGlnFrnPheserTyPrProtyGlu 508

RESULT 12
095714 PRELIMINARY: PRT; 4834 AA.
ID 095714.
AC 095714.
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE HERC2 protein.
GN HERC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99138701; PubMed=9949213;
RA Ji Y., Walkowicz M.J., Builing K., Johnson D.K., Tarvin R.E.,
RA Rinchik E.M., Horsthemke B., Stubbs L.J., Nicholls R.D.;
RT "The ancestral gene for transcribed, low-copy repeats in the Prader-
RT Willi/Angelman region encodes a large protein implicated in protein
RT trafficking, which is deficient in mice with neuromuscular and
RT spermatogenic abnormalities."
RL Hum. Mol. Genet. 8:533-542(1999).
CC -1- FUNCTION: MAY ACT AS A GUANINE NUCLEOTIDE EXCHANGE FACTOR INVOLVED
CC IN PROTEIN TRAFFICKING AND DEGRADATION PATHWAYS IN THE CELL. MAY
CC ALSO PLAY A ROLE IN UBIQUITIN CONJUGATION.
CC -1- SIMILARITY: CONTAINS 1 HECT-TYPE E3 UBIQUITIN-PROTEIN LIGASE
CC DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 RCC1-LIKE (REGULATOR OF CHROMOSOME
CC CONDENSATION-LIKE) DOMAINS.
DR EMBL; AF071172; AAD08657.1; -.
DR InterPro: IPR001199; Cyt_B5.
DR InterPro: IPR000569; HECT_domain.
DR InterPro: IPR000408; Reg_Chr_condens.
DR InterPro: IPR000433; Zn_Z2.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00173; heme_1; 1.
DR Pfam; PF00415; RCC1; 16.
DR Pfam; PF00569; Z2; 1.
DR PRINTS; PR00633; RCCONDSATON.
DR SMART; SM00119; HECTc; 1.

```


[illegible]

| | |
|----|--|
| RT | "A very large protein with diverse functional motifs is deficient in |
| RT | rtjs (runty, jerky, sterile) mice." |
| RL | Proc. Natl. Acad. Sci. U.S.A. 95:9436-9441(1998). |
| RL | [2] |
| RN | SEQUENCE FROM N.A. |
| RP | |
| RX | MEDLINE=99138701; PubMed=9949213; |
| RA | Ji Y., Walkowicz M.J., Bulting K., Johnson D.K., Taryn R.E., |
| RA | Ritschik E.M., Horsthemke B., Stubbs L.J., Nicholls R.D.; |
| RT | "The ancestral gene for transcribed, low-copy repeats in the Prader- |
| RT | Willi/Angelman region encodes a large protein implicated in protein |
| RT | trafficking, which is deficient in mice with neuromuscular and |
| RT | spermiogenic abnormalities"; |
| RL | Hum. Mol. Genet. 8:533-542(1999). |
| CC | -1- FUNCTION: MAY ACT AS A GUANINE NUCLEOTIDE EXCHANGE FACTOR INVOLVED |
| CC | IN PROTEIN TRAFFICKING AND DEGRADATION PATHWAYS IN THE CELL. MAY |
| CC | ALSO PLAY A ROLE IN UBIQUITIN CONJUGATION. |
| CC | -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BRAIN AND TESTIS WITH |
| CC | LOWER LEVELS IN HEART, LUNG, LIVER, SKELETAL MUSCLE AND KIDNEY. |
| CC | LITTLE EXPRESSION DETECTED IN SPLEEN. |
| CC | -1- DISEASE: DEFECTS IN HECT2 ARE THE CAUSE OF THE RUNTY JERKY STERILE |
| CC | (RJS) PHENOTYPE (ALSO KNOWN AS JUVENILE DEVELOPMENT AND FERTILITY |
| CC | (JDF)) WHICH IS CHARACTERIZED BY REDUCED GROWTH, JERKY GAIT, MALE |
| CC | STERILITY, FEMALE SEMISTERILITY AND MATERNAL BEHAVIOR DEFECTS. |
| CC | -1- SIMILARITY: CONTAINS 1 HECT-TYPE E3 UBIQUITIN-PROTEIN LIGASE |
| CC | DOMAIN. |
| CC | -1- SIMILARITY: CONTAINS 3 RCC1-LIKE (REGULATOR OF CHROMOSOME |
| CC | CONDENSATION-LIKE) DOMAINS. |
| DR | EMBL: AF061529; AAC31431.1; - |
| DR | EMBL: AF071173; AAD08658.1; - |
| DR | EMBL: AF071175; AAD08661.1; - |
| DR | EMBL: AF071177; AAD08659.1; - |
| DR | EMBL: AF071176; AAD08659.1; JOINED. |
| DR | MGI: 103234; Herc2. |
| DR | InterPro: IPR000345; Cyt_c_heme_bind. |
| DR | InterPro: IPR001199; Cyt_b5. |
| DR | InterPro: IPR000569; HECT_domain. |
| DR | InterPro: IPR000408; Reg_cht_condens. |
| DR | InterPro: IPR000433; ZnF_ZZ. |
| DR | Pfam: PF00632; HECT_1. |
| DR | Pfam: PF00173; heme_1; 1. |
| DR | Pfam: PF00415; RCC1_16. |
| DR | Pfam: PF00569; ZZ_1. |
| DR | PRINTS: PR00633; RCCNDNSATON. |
| DR | SMART: SM00119; HECTC; 1. |
| DR | SMART: SM00291; ZnF_ZZ; 1. |
| DR | PROSITE: PS50255; CYTOCHROME_B5_2; 1. |
| DR | PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1. |
| DR | PROSITE: PS50237; HECT_1. |
| DR | PROSITE: PS00625; RCC1_1; FALSE_NEG. |
| DR | PROSITE: PS00626; RCC1_2; UNKNOWN_2. |
| DR | PROSITE: PS50012; RCC1_3; 19. |
| DR | PROSITE: PS01357; ZZ_ZZ_1; UNKNOWN_1. |
| KV | Guanine-nucleotide releasing factor; Ubiquitin conjugation; Repeat. |
| FT | DOMAIN 515 780 RCC1-LIKE. |
| FT | DOMAIN 515 780 5 X TANDEM REPEATS. |
| FT | DOMAIN 515 570 1-1. |
| FT | REPEAT 571 622 1-2. |
| FT | REPEAT 623 676 1-2. |
| FT | REPEAT 677 728 1-4. |
| FT | REPEAT 729 780 1-5. |
| FT | DOMAIN 2560 3328 RCC1-LIKE. |
| FT | DOMAIN 2560 3328 7 X TANDEM REPEATS. |
| FT | REPEAT 3012 3066 2-1. |
| FT | REPEAT 3067 3118 2-2. |
| FT | REPEAT 3119 3170 2-3. |
| FT | REPEAT 3171 3224 2-4. |
| FT | REPEAT 3225 3276 2-5. |
| FT | REPEAT 3277 3328 2-6. |
| FT | DOMAIN 3954 4321 RCC1-LIKE. |
| FT | DOMAIN 3954 4321 7 X TANDEM REPEATS. |
| FT | REPEAT 4006 4059 3-1. |
| FT | REPEAT 4006 4059 3-2. |

FT REPEAT 4060 4111 3-3.
FT REPEAT 4112 4163 3-4.
FT REPEAT 4164 4217 3-5.
FT REPEAT 4218 4269 3-6.
FT REPEAT 4270 4321 3-7.
FT DOMAIN 4491 4796 HECT.
FT CONFLICT 692 724 L->V (IN REF. 2).
FT CONFLICT 724 747 R->G (IN REF. 2).
FT CONFLICT 747 756 L->F (IN REF. 2).
FT CONFLICT 756 929 R->P (IN REF. 2).
FT CONFLICT 929 1114 S->N (IN REF. 2).
FT CONFLICT 1114 1235 G->D (IN REF. 2).
FT CONFLICT 1235 1238 P->A (IN REF. 2).
FT CONFLICT 1238 2348 D->E (IN REF. 2).
FT CONFLICT 2348 2523 F->Y (IN REF. 2).
FT CONFLICT 2523 2567 L->V (IN REF. 2).
FT CONFLICT 2567 2572 E->Q (IN REF. 2).
FT CONFLICT 2572 3095 C->Y (IN REF. 2).
FT CONFLICT 3095 3107 A->P (IN REF. 2).
FT CONFLICT 3107 3114 T->S (IN REF. 2).
FT CONFLICT 3114 3161 LL->VV (IN REF. 2).
FT CONFLICT 3161 3386 A->V (IN REF. 2).
FT CONFLICT 3386 3508 M->I (IN AAD0861).
FT CONFLICT 3508 3712 S->C (IN REF. 2).
FT CONFLICT 3712 4187 S->T (IN REF. 2).
FT CONFLICT 4187 4716 R->C (IN REF. 2).
FT CONFLICT 4716 4723 R->C (IN REF. 2).
FT CONFLICT 4723 4730 S->R (IN REF. 2).
FT CONFLICT 4730 4752 Y->N (IN REF. 2).
FT CONFLICT 4752 4790 S->C (IN REF. 2).
FT CONFLICT 4790 4790 S->C (IN REF. 2).
SQ SEQUENCE 4836 AA; 527369 MM; 6A/604D77CD312AE CRC64;

Alignment Scores:

Pred. No.: 19.6 Length: 4836
Score: 50.50 Matches: 12
Percent Similarity: 60.71% Conservative: 5
Best Local Similarity: 42.86% Mismatches: 10
Query Match: 34.59% Indels: 1
DB: Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x 088473 (1-4836)

OY 83 CTATGCAATTTGTGAGAACACCGCTTCTTA--AAGTTAGCCACACAGAACACACAA 27
||||: ||| :||| |||: |||: |||
Db 4032 leuenguserileglnhisvalphelleylsvalaiaaasnerglyglylys 4051
OY 26 CAGCTGCTGCTGACGAGAGATGAA 3
Db 4052 Hiscysleualaleusergerln 4059
||||: ||| :|||

RESULT 14

09FHB4 PRELIMINARY; PRT; 241 AA.
AC 09FHB4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Modulin-like protein.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).

DR EMBL; AB019226; BAB10545.1; -
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR; 2
SQ SEQUENCE 241 AA; 27287 MM; 622935A377008A20 CRC64;

Alignment Scores:

Pred. No.: 23.2 Length: 241
Score: 50.00 Matches: 7
Percent Similarity: 75.00% Conservative: 5
Best Local Similarity: 43.75% Mismatches: 4
Query Match: 32.89% Indels: 0
DB: Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x 09FHB4 (1-241)

OY 31 TGTCTTCCTGTTGTGCTACTTAAACCGCGTTCACGAATTG 78
||||: ||| :||| |||: |||: |||
Db 167 Cysileasleuvalglyasleuglylsarlgcysheglulysile 182

RESULT 15

09VU37 PRELIMINARY; PRT; 299 AA.
AC 09VU37;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CG11268 protein (RH37735p).
CN CG11268.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayyan A., An H.-J., Andrews-Planck C., Baldwin D.,
RA Bailley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo G., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwu C.,
RA Jellali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobaraj B., Morris J., Moshirei A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Spierkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RA Stapleton M., Brokstein P., Hong L., Abhayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Ceiniker S.,

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003539; AAF49854.1; -

DR EMBL; AY071003; AAL48625.1; -

DR EMBL; AY070700; AAL48171.1; -

DR FlyBase; FBgn0036336; CG11268.

SO SEQUENCE 299 AA; 34104 MW; 619350D9416E71AF CRC64;

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-----|
| Pred. No.: | 23.2 | Length: | 299 |
| Score: | 50.00 | Matches: | 9 |
| Percent Similarity: | 66.67% | Conservative: | 5 |
| Best Local Similarity: | 42.86% | Mismatches: | 7 |
| Query Match: | 34.25% | Indels: | 0 |
| DB: | 5 | Gaps: | 0 |

US-09-939-293-1_COPY_56_139 (1-84) x Q9VU37 (1-299)

QY 83 CTTATCAATTCGTGAGAAACACCGCTTCTTAAGTTCACACACAGACACACACAAACAC 24

Db 204 LeuValGlnAsnGluLysHisSerGluHisLysLeuIleThrSerGlyIleTyrAlaTyr 223

QY 23 TGT 21

Db 224 Cys 224

Search completed: February 16, 2003, 03:10:23
 Job time : 27.7612 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: February 16, 2003, 01:22:19 : Search time 7.99254 Seconds
(without alignments)
871.816 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 152

Sequence: 1 actctctctcagctacagc.....gtttcctcagatgtataga 84

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09939293/runat.12022003_170353_8513/app_query.fasta.1.590
-DB=SwissProt_40 -OPT=fastcan -SUFFIX=isp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09939293 @CGN.1.1.10 @runat.12022003_170353_8513 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TTIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 152 | 100.0 | 239 | 1 | SMAC_HUMAN |
| 2 | 87 | 57.2 | 237 | 1 | SMAC_MOUSE |
| 3 | 51 | 34.9 | 2594 | 1 | 7LBS_DROVI |
| 4 | 50 | 32.9 | 432 | 1 | PARI_RAT |
| 5 | 49 | 32.2 | 366 | 1 | ML1A_SHEEP |
| 6 | 49 | 32.2 | 1391 | 1 | RPCL_HUMAN |
| 7 | 49 | 32.2 | 3924 | 1 | ANK2_HUMAN |
| 8 | 48.5 | 33.2 | 773 | 1 | CPT1_HUMAN |
| 9 | 48 | 32.9 | 763 | 1 | ECHA_RAT |
| 10 | 47 | 30.9 | 593 | 1 | STB2_CANFA |
| 11 | 47 | 30.9 | 593 | 1 | STB2_HUMAN |
| 12 | 47 | 30.9 | 594 | 1 | STB2_MOUSE |
| 13 | 47 | 30.9 | 594 | 1 | STB2_RAT |
| 14 | 47 | 32.2 | 749 | 1 | NU5C_ARBJA |
| 15 | 47 | 30.9 | 1530 | 1 | RPCL_TRYBB |
| 16 | 47 | 30.9 | 1912 | 1 | PTPD_HUMAN |
| 17 | 47 | 30.9 | 2359 | 1 | RPCL_PLAFA |
| 18 | 47 | 30.9 | 4377 | 1 | ANK3_HUMAN |

| | | | | | | | |
|------|------|------|------|---|------------|--------|--------------|
| C 19 | 46.5 | 31.8 | 764 | 1 | CPT1_MOUSE | P97742 | mus musc |
| C 20 | 46.5 | 31.8 | 772 | 1 | CPT1_RAT | O63704 | rattus norv |
| C 21 | 46.5 | 31.8 | 773 | 1 | CPT1_RAT | P32198 | rattus norv |
| C 22 | 46 | 31.5 | 297 | 1 | PTR_METH | P21348 | methanobact |
| C 23 | 46 | 31.5 | 504 | 1 | C6B2_HELM | Q27684 | helicoverpa |
| C 24 | 46 | 30.3 | 522 | 1 | SOA2_HUMAN | O75908 | homo sapien |
| C 25 | 46 | 30.3 | 525 | 1 | SOA2_MOUSE | O88908 | mus musc |
| C 26 | 46 | 30.3 | 526 | 1 | SOA2_CERE | O77759 | ceropithec |
| C 27 | 46 | 31.5 | 606 | 1 | RA17_SCPO | P50531 | schistosach |
| C 28 | 46 | 31.5 | 690 | 1 | SYGB_BUCAT | P57235 | buchnera ap |
| C 29 | 45.5 | 29.9 | 110 | 1 | VNBP_HELV | Q00572 | helicobact |
| C 30 | 45.5 | 29.9 | 520 | 1 | CC3_YEAST | P32457 | saccharomyc |
| C 31 | 45.5 | 31.2 | 588 | 1 | Y06B_CAEEL | P34602 | caenorhabd |
| C 32 | 45.5 | 29.9 | 668 | 1 | UVRC_IACEL | O9c688 | lactococcu |
| C 33 | 45.5 | 29.9 | 1200 | 1 | ICEN_PSRXY | P06650 | pseudomonas |
| C 34 | 45 | 29.6 | 230 | 1 | FRHG_MERJA | O60340 | methanococ |
| C 35 | 45 | 30.8 | 372 | 1 | NTF3_TOBAC | Q40517 | nicotiana t |
| C 36 | 45 | 30.8 | 384 | 1 | MAPK_PETHY | Q40864 | petunia hyb |
| C 37 | 45 | 30.8 | 623 | 1 | SKN1_CAEEL | P34707 | caenorhabd |
| C 38 | 45 | 30.8 | 738 | 1 | NU5C_IACSA | O32539 | lactuca sat |
| C 39 | 45 | 30.8 | 741 | 1 | NU5C_CICIN | O32007 | cthorium i |
| C 40 | 45 | 30.8 | 743 | 1 | NU5C_CARTI | O32051 | cartilagus t |
| C 41 | 45 | 30.8 | 1442 | 1 | DP03_UREPA | O3p9d4 | ureaplasma |
| C 42 | 45 | 29.6 | 1460 | 1 | RPCL_YEAST | P04051 | saccharomyc |
| C 43 | 45 | 29.6 | 2911 | 1 | FBN2_HUMAN | P35556 | homo sapien |
| C 44 | 44.5 | 30.5 | 200 | 1 | FIMB_ECOLI | P04742 | escherichia |
| C 45 | 44.5 | 30.5 | 262 | 1 | TRPC_AZOB | P26938 | azospirillum |

ALIGNMENTS

RESULT 1

| | | | | |
|----|--|-----------|------|---------|
| ID | SMAC_HUMAN | STANDARD: | PRT: | 239 AA. |
| AC | Q9NR28: Q9NR28; Q9H4V6: Q96LV0; | | | |
| DT | 16-OCT-2001 (Rel. 40, Created) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update) | | | |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) | | | |
| DE | Smac protein, mitochondrial precursor (Second mitochondrial derived activator of caspase) (Direct IAP binding protein with low pI). | | | |
| GN | SMAC OR DIABLO. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1), PARTIAL SEQUENCE, FUNCTION, AND TISSUE SPECIFICITY. | | | |
| RX | MEDLINE=20383536; PubMed=10929711; | | | |
| RA | Du C., Fang M., Li Y., Li L., Wang X.; | | | |
| RT | "Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition."; | | | |
| RL | Cell 102:33-42(2000). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1). | | | |
| RA | Matanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T., Suzuki Y., Okeyashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.; | | | |
| RT | "MEDO human cDNA sequencing project."; | | | |
| RL | Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases. | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION. | | | |
| RA | Pubmed=10950947; | | | |
| RA | Stinivasula S.M., Datta P., Fan X.J., Fernandes-Alnemri T., Huang Z., Alnemri E.S.; | | | |
| RT | "Molecular determinants of the caspase-promoting activity of Smac/Diablo and its role in the death receptor pathway."; | | | |
| RL | J. Biol. Chem. 275:36152-36157(2000). | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. (ISOFORM 1). | | | |
| RA | Tissue=Cerebellum; | | | |
| RA | Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y., Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M., | | | |

RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
 RA Kawakami B., Nagai K., Isogai T., Sugano S.,
 RT "NEO human cDNA sequencing project."
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RL [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Muscle, and Uterus;
 RL Strassberg R.;
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 RL [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 56-239.
 RX MEDLINE-20426096; PubMed-10972280;
 RA Chal J., Du C., Wu J.W., Kyin S., Wang X., Shi Y.,
 RT "Structural and biochemical basis of apoptotic activation by
 RT Smac/DIABLO."
 RL Nature 406:855-862(2000).
 RL [7]
 RP STRUCTURE BY NMR OF 56-64 IN COMPLEX WITH BIRC4.
 RX MEDLINE-21020961; PubMed-11140637;
 RA Liu Z., Sun C., Olejniczak E.T., Meadows R.P., Betz S.F., Coost T.,
 RA Hermann J., Wu J.C., Pesik S.W.;
 RT "Structural basis for binding of Smac/DIABLO to the XIAP BIR3
 RT domain."
 RL Nature 408:1004-1008(2000).
 CC -I- FUNCTION: PROMOTES APOPTOSIS BY ACTIVATING CASPASES IN THE
 CC CYTOCHROME C/PAF-1/CASPASE-9 PATHWAY. ACTS BY OPPOSING THE
 CC INHIBITORY ACTIVITY OF INHIBITOR OF APOPTOSIS PROTEINS (IAP).
 CC -I- SUBUNIT: Homodimer. Interacts with BIRC2, BIRC3, BIRC4/XIAP and
 CC BIRC7.
 CC -I- SUBCELLULAR LOCATION: MITOCHONDRIAL BUT RELEASED INTO THE CYTOSOL
 CC WHEN CELLS UNDERGO APOPTOSIS.
 CC -I- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2/Diablo-S;
 CC are produced by alternative splicing.
 CC -I- TISSUE SPECIFICITY: UNBOUTIOUSLY EXPRESSED WITH HIGHEST EXPRESSION
 CC IN TESTIS. EXPRESSION IS ALSO HIGH IN HEART, LIVER, KIDNEY,
 CC SPLEEN, PROSTATE AND OVARY. LOW IN BRAIN, LUNG, THYMUS AND
 CC PERIPHERAL BLOOD LEUKOCYTES.
 CC -I- DOMAIN: The mature N-terminus mediates interaction with
 CC BIRC4/XIAP.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF262240; AAF87716.1; -;
 DR EMBL: AK024768; BAB14994.1; -;
 DR EMBL: AF298770; AAG22077.1; -;
 DR EMBL: AK057778; BAB71568.1; -;
 DR EMBL: BC004417; AAH04417.1; -;
 DR PDB: 1FEW; 13-SEP-00.
 DR PDB: 1G3F; 10-JAN-01.
 DR MIM: 605219; -;
 KM Transit peptide; Mitochondrion; Apoptosis; Alternative splicing;
 KM 3D-structure.
 FT TRANSIT 1 55 MITOCHONDRION.
 FT CHAIN 56 239 IAP-BINDING MOTIF (BY SIMILARITY).
 FT SITE 56 60 MAALSKMSFRYSVTSFRQCLCPVYANFKKRCFSELIRP
 FT VARSPLIC 1 60 MKRWTVIGWVTLCAVPPIA -> MKSDPYE (IN
 FT ISOFORM 2).
 FT CONFLICT 32 32 K -> E (IN REF. 4).
 FT CONFLICT 44 44 K -> R (IN REF. 2).
 FT CONFLICT 62 105 MISSING (IN REF. 4).
 FT CONFLICT 165 165 E -> K (IN REF. 4).
 SQ SEQUENCE 239 AA; 27131 MW; 70C2AEDDC654D031 CRC64;
 Alignment Scores: 2,68e-16 Length: 239
 Pred. No.: 152.00 Matches: 28

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0
 US-09-939-293-1_COPY_56_139 (1-84) x SMAC_HUMAN (1-239)
 QY 1 ACTTCATTCCTTCAGGTACGACAGTCTTGTCTGCTGTGGCTAATTGAAG 60
 DB 13 ThSRPhePhenArgYrArgIncysIleuGysValProValAlaIAsnPhenyls 32
 QY 61 CGGTGTTTCTCAGATTGATAGA 84
 DB 33 ArgCysPheSerGluIleuLearg 40
 RESULT 2
 ID SMAC_MOUSE STANDARD: PRT; 237 AA.
 AC Q9JIO3; Q9CZD1; Q9DCD3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Smac protein, mitochondrial precursor (Second mitochondria-derived
 DE activator of caspase) (Direct IAP binding protein with low pI).
 GN SMAC OR DIABLO.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
 RP SPECIFICITY.
 RC STRAIN=BALB/c; TISSUE=Kidney;
 RX MEDLINE-20383537; PubMed-10929712;
 RA Verhaegen A.M., Ekert P.G., Pakusch M., Silke J., Connolly L.M.,
 RA Reid G.E., Moritz R.L., Simpson R.J., Vaux D.L.;
 RT "Identification of DIABLO, a mammalian protein that promotes apoptosis
 RT by binding to and antagonizing IAP proteins."
 RL Cell 102:43-53(2000).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE-21085660; PubMed-11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
 RA Kuenl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Bash G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombeerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whitlaker C.,
 RA Wilming L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H.,
 RA Kohetsuki S.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -I- FUNCTION: PROMOTES APOPTOSIS BY ACTIVATING CASPASES IN THE
 CC CYTOCHROME C/PAF-1/CASPASE-9 PATHWAY. ACTS BY OPPOSING THE
 CC INHIBITORY ACTIVITY OF INHIBITOR OF APOPTOSIS PROTEINS (IAP).
 CC -I- SUBUNIT: Homodimer. Interacts with BIRC2, BIRC3, BIRC4/XIAP and
 CC BIRC7 (By similarity).
 CC -I- SUBCELLULAR LOCATION: MITOCHONDRIAL BUT RELEASED INTO THE CYTOSOL
 CC WHEN CELLS UNDERGO APOPTOSIS.
 CC -I- TISSUE SPECIFICITY: HIGHEST EXPRESSION FOUND IN HEART, LIVER,
 CC KIDNEY AND TESTIS.
 CC -I- DOMAIN: The mature N-terminus mediates interaction with

CC BIRC4/XIAP (By similarity).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF203914; AAF82190.1; -
DR EMBL: AK012760; BAB28450.1; -
DR EMBL: AK002887; BAB22433.1; ALF_FRAME.
DR HSSP: Q9NR28; 1FEW.
DR MGD: MG1:1913843; 0610041G12R1k.
KW Transit peptide; Mitochondrion; Apoptosis.
FT TRANSIT 1 53 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 54 237 SMAC PROTEIN.
FT SITE 54 58 IAP-BINDING MOTIF (BY SIMILARITY).
FT CONFLICT 64 64 H -> Q (IN REF. 2).
SQ SEQUENCE 237 AA; 26829 MW; E53B6F04FC390A1 CRC64;

Alignment Scores:
Pred. No.: 7.51e-06 Length: 237
Score: 87.00 Matches: 19
Percent Similarity: 77.78% Conservative: 2
Best Local Similarity: 70.37% Mismatches: 4
Query Match: 57.24% Indels: 2
DB: Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x SMAC_MOUSE (1-237)

QY 4 TCATTTCTTCAGTACACAGCTGTTTGTGCTTGTGCTTACTTGAACGCG 63
DB 14 SerleupheargtyrAagGlnAargpHe-----ProvalleuAlaaseneLysLysArg 31

QY 64 TGTTCCTCAGATTGATAGAGA 84
DB 32 Cyspheserluleuilellys 38

RESULT 3
ID 7LES_DROVI STANDARD; PRT; 2594 AA.
AC P20806:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sevenless protein (EC 2.7.1.112).
GN SEV.
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7244;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90319110; PubMed-2115169;
RA Michael W.M., Bowtell D.D.L., Rubin G.M.;
RT "Comparison of the sevenless genes of Drosophila virilis and
RT Drosophila melanogaster";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5351-5353(1990).
CC -1- FUNCTION: RECEPTOR FOR AN EXTRACELLULAR SIGNAL. REQUIRED TO
CC INSTRUCT A CELL TO DIFFERENTIATE INTO A R7 PHOTORECEPTOR. THE
CC LIGAND FOR SEV IS THE BOSS (BRIDE OF SEVENLESS) PROTEIN.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC tyrosine phosphate.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
CC RECEPTOR SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 7 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- CAUTION: UNCLEAR WHETHER THE POTENTIAL MEMBRANE SPANNING REGION
CC NEAR THE N-TERMINUS IS PRESENT AS A TRANSMEMBRANE DOMAIN IN THE
CC NATIVE PROTEIN OR SERVES AS A CLEAVED SIGNAL SEQUENCE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M34545; AAA28883.1; -
DR EMBL: M34544; AAA28883.1; JOINED.
DR EMBL: M34543; AAA28883.1; JOINED.
DR PIR: A35774; A35774.
DR HSSP: P11362; 1FGK.
DR FLYBASE: FBgn0013140; Dv1r/sev.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR000033; Ldl_receptor_rep.
DR InterPro: IPR002011; RtkkinaseII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00041; fn3; 6.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00060; FN3; 5.
DR SMART: SM00135; LY; 1.
DR SMART: SM00219; TYRKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Tyrosine-protein kinase; Receptor; Vision; Transmembrane;
KW Glycoprotein; ATP-binding; Phosphorylation; Repeat.
FT DOMAIN 1 2139 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 1 2140 POTENTIAL.
FT DOMAIN 2164 2594 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 9 26 POLY-GLN.
FT DOMAIN 356 459 FIBRONECTIN TYPE-III 1.
FT DOMAIN 464 555 FIBRONECTIN TYPE-III 2.
FT DOMAIN 835 935 FIBRONECTIN TYPE-III 3.
FT DOMAIN 1328 1421 FIBRONECTIN TYPE-III 4.
FT DOMAIN 1706 1816 FIBRONECTIN TYPE-III 5.
FT DOMAIN 1817 1916 FIBRONECTIN TYPE-III 6.
FT DOMAIN 1917 2007 FIBRONECTIN TYPE-III 7.
FT DOMAIN 2057 2063 POLY-ARG.
FT DOMAIN 2224 2495 PROTEIN KINASE.
FT NP_BIND 2230 2238 ATP (BY SIMILARITY).
FT BINDING 2257 2257 ATP (BY SIMILARITY).
FT MOD_RES 2391 2391 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 77 77 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 508 508 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 641 641 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 778 778 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 797 797 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 874 874 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1257 1257 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1344 1344 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1382 1382 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1577 1577 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1587 1587 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1665 1665 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1752 1752 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1776 1776 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1824 1824 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1908 1908 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1966 1966 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 2088 2088 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 2594 AA; 289130 MW; 77D8A356CBAD0BDB CRC64;

Alignment Scores:
Pred. No.: 5.4 Length: 2594

Score: 51.00 Matches: 9
 Percent Similarity: 69.57% Conservative: 7
 Best Local Similarity: 39.13% Mismatches: 7
 Query Match: 34.93% Indels: 0
 DB: 1 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x 7LES_DPROVI (1-2594)

QY 71 GAGAAACACCGCTCTTAAGTACGACACAGACACAAACAGCTGTGACCTG 12

Db 757 GAGAAACACCGCTCTTAAGTACGACACAGACACAAACAGCTGTGACCTG 776

QY 11 AAGATGAA 3

Db 777 SerAsnAsp 779

RESULT 4

PAR1_RAT STANDARD; PRT; 432 AA.

AC P26824;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).

GN F2R OR PAR1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

NCBI_TaxID=10116;

RP SEQUENCE FROM N.A.

RC TISSUE=Arteric smooth muscle;

RA MEDLINE=92381002; PubMed=1324917;

RT "Molecular cloning of the rat vascular smooth muscle thrombin receptor. Evidence for in vitro regulation by basic fibroblast growth factor.";

RT J. Biol. Chem. 267:16975-16979(1992).

CC -I- FUNCTION: High affinity receptor for activated thrombin coupled to G proteins that stimulate phosphoinositide hydrolysis.

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- PTM: A proteolytic cleavage generates a new amino terminus that functions as a tethered ligand.

CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SMSS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M81642; AAA42274.1; -

DR PIR; A43448; A43448.

DR HSSP; P34996; 1DD.

DR InterPro; IPR00276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm1.1; 1.

DR PRINTS; PR00237; GPCR_Rhodpsn.

DR PROSITE; PS00237; G-PROTEIN_RECP_FL_1; 1.

DR PROSITE; PS50262; G-PROTEIN_RECP_FL_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Blood coagulation.

FT SIGNAL 1 21 POTENTIAL.

FT PROPEP 22 45 REMOVED FOR RECEPTOR ACTIVATION (BY SIMILARITY).

FT CHAIN 46 432 PROTEINASE ACTIVATED RECEPTOR 1.

FT DOMAIN 46 109 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 110 135 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 136 144 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 145 164 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 165 183 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 184 205 3 (POTENTIAL).

FT DOMAIN 206 225 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 226 246 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 247 275 5 (POTENTIAL).

FT TRANSMEM 276 295 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 296 318 6 (POTENTIAL).

FT TRANSMEM 319 341 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 342 357 7 (POTENTIAL).

FT TRANSMEM 358 381 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 382 432 ASP/GLU-RICH (ACIDIC).

FT DOMAIN 433 482 POLY-PRO.

FT SITE 45 46 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).

FT DISULFID 182 261 BY SIMILARITY.

FT CARBOHYD 69 82 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 82 86 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 257 266 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 266 266 N-LINKED (GLCNAc. . .) (POTENTIAL).

SQ SEQUENCE 432 AA; 48280 MW; DD032B97ABA4A605 CRC64;

Alignment Scores:

Pred. No.:

Score: 6.93

Percent Similarity: 50.00

Best Local Similarity: 55.56%

Query Match: 40.74%

DB: 1 Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x PAR1_RAT (1-432)

QY 1 ACTTCATCTTTCAGGTACAGACAGTGTGTGCTGTGCTGCTTAAGTAAAG 60

Db 355 ThrAlaTyrPheAlaTyrLeuLeuLeuValCysValThrSerValAlaSer 371

QY 61 CGGTGTTCTCAGATTCATA 81

Db 372 ---CysIleAspProLeuIle 377

RESULT 5

MLA_SHEEP STANDARD; PRT; 366 AA.

AC P48040; 046608;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Melatonin receptor type 1A (Mel-1A-R).

GN MTNRLA.

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

NCBI_TaxID=9940;

RP SEQUENCE FROM N.A.

RC TISSUE=pituitary pars tuberalis;

RA MEDLINE=95033233; PubMed=7946354;

RA Reppert S.M., Weaver D.R., Ebisawa T.;

RT "Cloning and characterization of a mammalian melatonin receptor that mediates reproductive and circadian responses.";

RT Neuron 13:1177-1185(1994).

CC -----

DR MEDLINE=97337902; PubMed=9194573;

DR Barrett P., Conway S., Jockers R., Strosberg A.D.,

RA Guardiola-Iemaitre B., Delagrègne P., Morgan P.J.;

RT "Cloning and functional analysis of a polymorphic variant of the ovine Mel 1a melatonin receptor.";

RT Biochim. Biophys. Acta 1356:299-307(1997).

CC -I- FUNCTION: HIGH AFFINITY RECEPTOR FOR MELANOTIN. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY PERUSSIS TOXIN SENSITIVE G

CC PROTEINS THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.

CC -I- SUBCELLULAR LOCATION: Integral membrane protein.

CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.


```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U14109; AAB17721.1; -
DR EMBL: AF045219; AAC02699.1; -
DR HSP: P29274; 1MMH.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm1.1; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
FT DOMAIN 1 45
FT TRANSMEM 46 66
FT TRANSMEM 67 79
FT TRANSMEM 80 100
FT TRANSMEM 101 118
FT TRANSMEM 119 139
FT TRANSMEM 140 158
FT TRANSMEM 159 179
FT TRANSMEM 180 203
FT TRANSMEM 204 224
FT TRANSMEM 225 236
FT TRANSMEM 237 277
FT DOMAIN 278 290
FT TRANSMEM 291 311
FT TRANSMEM 312 366
FT DISULFID 116 193
FT CARBOHYD 16 16
FT CARBOHYD 23 23
FT VARIANT 282 282 A -> D (IN MEL 1A(BETA)).
FT VARIANT 358 358 H -> R (IN MEL 1A(BETA)).
FT VARIANT 361 361 I -> V (IN MEL 1A(BETA)).
SQ SEQUENCE 366 AA: 40400 MW: 5386EDDF9710E4A CRC64;

Alignment Scores:
Pred. No.: 9.93 Length: 366
Score: 49.00 Matches: 12
Percent Similarity: 88.24% Conservative: 3
Best Local Similarity: 70.59% Mismatches: 0
Query Match: 32.24% Indels: 2
DB: 1 Gaps: 2

US-09-939-293-1_COPY_56_139 (1-84) x ML1A_SHEEP (1-366)

OY 11 TCAGGTACAGACAGTGTGTGT---GTGTCCTGCTGTGTG---CTACTTTA 55
DB 156 SerGlyThrAsnSerLeuGlyValPheLeuIleTrpThrLeuThrLeu 172

RESULT 6
RPL1_HUMAN STANDARD: PRT: 1391 AA.
AC 014802;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DE 16-OCT-2001 (rel. 40, Last annotation update)
DE DNA-directed RNA polymerase III largest subunit (EC 2.7.7.6) (RPL155).
GN POLR3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=97474795; PubMed=9331371;
RA Sepehri S.; Hernandez N.;
RT "The largest subunit of human RNA polymerase III is closely related
RT to the largest subunit of yeast and trypanosome RNA polymerase III.";
```

```
RL Genome Res. 7:1006-1019(1997).
CC -1 FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1 CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1 SUBUNIT: RNA POLYMERASE III CONSISTS OF ABOUT 15 DIFFERENT
CC SUBUNITS. THIS SUBUNIT IS THE LARGEST COMPONENT OF RNA POLYMERASE
CC III.
CC -1 SUBCELLULAR LOCATION: Nuclear.
CC -1 MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND 5.8S GENES.
CC -1 SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF021351; AAB86536.1; -
DR InterPro: IPR000722; RNA_pol_A.
DR InterPro: IPR002879; RNA_pol_A2.
DR Pfam: PF0623; RNA_pol_A; 1.
DR Pfam: PF01854; RNA_pol_A2; 1.
KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger; Nuclear protein.
FT ZN_FING 69 82
FT SEQUENCE 1391 AA: 155748 MW: 500BE9C21F32387C CRC64;

Alignment Scores:
Pred. No.: 10.9 Length: 1391
Score: 49.00 Matches: 7
Percent Similarity: 73.33% Conservative: 4
Best Local Similarity: 46.67% Mismatches: 4
Query Match: 32.24% Indels: 0
DB: 1 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x RPL1_HUMAN (1-1391)

OY 10 TTCAGTACAGACAGTGTGTGTGTGTCCTGCTGTGTGCGTACTTT 54
DB 486 PheArgPheAsnGlnGlyValCysTrpProTyrAsnAlaAspPhe 500

RESULT 7
ANK2_HUMAN STANDARD: PRT: 3924 AA.
AC 001484; 001485;
DT 01-APR-1993 (rel. 25, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=9130246; PubMed=1830053;
RA Otto E.; Kunimoto M.; McLaughlin T.; Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
RT ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).
RP [2]
RP REVISIONS.
RA Carpenter S.;
RT submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
```



```

RX MEDLINE=95199277; PubMed=7892212;
RA Britton C.H., Schultz R.A., Zhang B., Esser V., Foster D.W.,
RA McGarrity J.D.;
RT "Human liver mitochondrial carnitine palmitoyltransferase I:
RT characterization of its cDNA and chromosomal localization and partial
RT analysis of the gene."
RL Proc. Natl. Acad. Sci. U.S.A. 92:1984-1988(1995).
CC -1- CATALYTIC ACTIVITY: Palmitoyl-CoA + L-carnitine = CoA + L-
CC palmitoylcarnitine.
CC -1- ENZYME REGULATION: INHIBITORS SUCH AS MALONYL-COA INTERACT WITH
CC ITS CATALYTIC DOMAIN AND NOT WITH AN ASSOCIATED REGULATORY
CC COMPONENT.
CC -1- PATHWAY: Fatty acid beta-oxidation cycle.
CC -1- SUBCELLULAR LOCATION: Mitochondrial outer membrane.
CC -1- TISSUE SPECIFICITY: STRONG EXPRESSION IN KIDNEY AND HEART, AND
CC LOWER IN LIVER AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE CARNITINE/CHOLINE ACETYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L39211; AAC41748.1; -.
DR Genew: HGNC:2328; CPT1A.
DR MIM: 600528; -.
DR DR MIM: 255120; -.
DR InterPro: IPR000542; Carn.acyltransf.
DR Pfam: PF00755; Carn.acyltransf; 1.
DR PROSITE: PS00438; ACYLTRANSF_C_1; 1.
DR PROSITE: PS00440; ACYLTRANSF_C_2; 1.
DR Transfaser: Acyltransferase; Mitochondrion; Outer membrane;
DR Fatty acid metabolism; Transport; Transmembrane; Multigene family.
FT DOMAIN 1 47 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 48 73 POTENTIAL.
FT DOMAIN 74 102 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT TRANSSEM 103 122 POTENTIAL.
FT DOMAIN 123 773 CYTOPLASMIC (POTENTIAL).
FT ACT_SITE 473 473 POTENTIAL.
SQ SEQUENCE 773 AA; 88428 MW; 44639E228343C990 CRC64;

Alignment Scores:
Pred. No.: 12.6 Length: 773
Score: 48.50 Matches: 12
Percent Similarity: 65.38% Conservative: 5
Best Local Similarity: 46.15% Mismatches: 4
Query Match: 33.22% Indels: 5
DB: 1 Gaps: 2

US-09-939-293-1_COPY_56_139 (1-84) x CPT1_HUMAN (1-773)
QY 74 TCTGAGAAACGCGCTCTTA--AGTTAGCACACAGAGACA-----CAC 30
DB 637 SerIuIySHISGlnHISMeTTrArgLeuAlaMeTherGlyIleAspArgHis 656
QY 29 AAACACTGTCTGTACTG 12
DB 657 LeupheCysLeuTYrVal 662

RESULT 9
ECHA_RAT STANDARD; PRT; 763 AA.
AC 064428;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-JUN-2002 (Rel. 41, Last annotation update)
DE Trifunctional enzyme alpha subunit, mitochondrial precursor (pP-alpha)
DE [Includes: Long-chain enoyl-CoA hydratase (EC 4.2.1.17); Long chain
DE 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)].

```

```

GN HADHA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRATN-Wistar;
RX MEDLINE=94075334; PubMed=8253773;
RA Kamijo T., Aoyama T., Miyazaki J., Hashimoto T.;
RT "Molecular cloning of the cDNAs for the subunits of rat mitochondrial
RT fatty acid beta-oxidation multienzyme complex. Structural and
RT functional relationships to other mitochondrial and peroxisomal beta-
RT oxidation enzymes."
RL J. Biol. Chem. 268:26452-26460(1993).
CC -1- FUNCTION: BIFUNCTIONAL SUBUNIT.
CC -1- CATALYTIC ACTIVITY: (3S)-3-hydroxyacyl-CoA = trans-2(or 3)-enoyl-
CC CoA + H(2)O.
CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxyacyl-CoA + NAD(+) = 3-oxoacyl-CoA
CC + NADH.
CC -1- PATHWAY: CATALYZES TWO STEPS OF THE FATTY ACID BETA-OXIDATION
CC CYCLE.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE ENOYL-COA
CC HYDRATASE/ISOMERASE FAMILY.
CC -1- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE 3-
CC HYDROXYACYL-COA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D16478; BAA03939.1; -.
DR HSSP: P00348; 3BDH.
DR InterPro: IPR002135; 3HCDH.
DR InterPro: IPR001753; EnCoA_hydrase.
DR Pfam: PF00378; ECH; 1.
DR Pfam: PF00725; 3HCDH; 1.
DR Pfam: PF02737; 3HCDH_N; 1.
DR PROSITE: PS00067; 3HCDH; 1.
DR PROSITE: PS00166; ENOYL_COA_HYDRATASE; 1.
DR Fatty acid metabolism; Multifunctional enzyme; Oxidoreductase; NAD;
DR Lyase; Mitochondrion; Transit peptide.
FT TRANSIT 1 36 MITOCHONDRION (POTENTIAL).
FT CHAIN 37 763 TRIFUNCTIONAL ENZYME ALPHA SUBUNIT.
FT ACT_SITE 151 151 ACTIVATES A WATER MOLECULE (BY
FT ACT_SITE 173 173 SIMILARITY).
FT ACT_SITE PROVIDES THE ALPHA-PROTON (BY
FT ACT_SITE SIMILARITY).
SQ SEQUENCE 763 AA; 82512 MW; F4A489C360AFA6 CRC64;

Alignment Scores:
Pred. No.: 15.1 Length: 763
Score: 48.00 Matches: 8
Percent Similarity: 61.90% Conservative: 5
Best Local Similarity: 38.10% Mismatches: 8
Query Match: 32.88% Indels: 0
DB: 1 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x ECHA_RAT (1-763)
QY 68 AAACACGCGCTCTTAAGTTCAGCACACAGACAGAACTGTCTACTCTGAAG 9
DB 455 LysHISLysValLeuLysGlnValGlnSerValThrProGlnHISGlyIlePheAlaSer 474
QY 8 AAT 6
DB 475 Asn 475

```

```

RESULT 10
RA STB2_CANFA STANDARD: PRT: 593 AA.
AC 028288:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
GN Syntaxin binding protein 2 (Unc-18 homolog 2) (Unc-18B) (Unc18-2).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96370811; PubMed=8774707;
RA Riento K., Jaentti J., Jansson S., Hielm S., Lehtonen E., Ehnholm C.,
RA Keranen S., Oikonen V.M.;
RT "A secret-related vesicle-transport protein that is expressed
RT predominantly in epithelial cells."
RL Eur. J. Biochem. 239:638-646(1996).
CC -I- FUNCTION: INVOLVED IN THE PROTEIN TRAFFICKING FROM THE GOLGI
CC APPARATUS TO THE PLASMA MEMBRANE. BINDS SYNTAXINS 1A, 2, 3 BUT NOT
CC 4 (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE STXBP/UNC-18/SECI FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL, LA1609; AAC00031.1;
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1.
KM Protein transport.
SO SEQUENCE 593 AA; 66499 MW; BC42120B638FC7E7 CRC64;

Alignment Scores:
Pred. No.: 21.5 Length: 593
Score: 47.00 Matches: 9
Percent Similarity: 77.78% Conservative: 5
Best Local Similarity: 50.00% Mismatches: 4
Query Match: 30.92% Indels: 0
DB: 1 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x STB2_CANFA (1-593)
OY 20 GACAGGTTTGTGTCCTCGTGTGGCTACTTAAGAAGCGGTGTTCTCAG 73
DB 402 AsplysillearyGalleuleuleuTyrlleuleuLargnsnglyValserGlu 419

RESULT 11
STB2_HUMAN STANDARD: PRT: 593 AA.
AC 015833:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
GN Syntaxin binding protein 2 (Unc-18 homolog 2) (Unc-18B) (Unc18-2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97079654; PubMed=8921365;
RA Ziegler S.F., Mortrud M.T., Swartz A.R., Baker E., Sutherland G.R.,

```

```

RA Burmeister M., Mulligan J.T.;
RT "Molecular characterization of a nonneuronal human UNC18 homolog.";
RL Genomics 37:19-23(1996).
CC -I- FUNCTION: INVOLVED IN THE PROTEIN TRAFFICKING FROM THE GOLGI
CC APPARATUS TO THE PLASMA MEMBRANE. BINDS SYNTAXINS 1A, 2, 3 BUT NOT
CC 4 (BY SIMILARITY).
CC -I- TISSUE SPECIFICITY: PLACENTA, LUNG, LIVER, KIDNEY AND PANCREAS, AS
CC WELL AS IN PERIPHERAL BLOOD LYMPHOCYTES.
CC -I- SIMILARITY: BELONGS TO THE STXBP/UNC-18/SECI FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL, U63533; AAC50762.1;
DR Genew; HGNC:11445; STXBP2.
DR MIM: 601717;
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1.
KM Protein transport.
SO SEQUENCE 593 AA; 66438 MW; 5BE27B47B2017A2B CRC64;

Alignment Scores:
Pred. No.: 21.5 Length: 593
Score: 47.00 Matches: 9
Percent Similarity: 77.78% Conservative: 5
Best Local Similarity: 50.00% Mismatches: 4
Query Match: 30.92% Indels: 0
DB: 1 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x STB2_HUMAN (1-593)
OY 20 GACAGGTTTGTGTCCTCGTGTGGCTACTTAAGAAGCGGTGTTCTCAG 73
DB 402 AsplysillearyGalleuleuleuTyrlleuleuLargnsnglyValserGlu 419

RESULT 12
STB2_MOUSE STANDARD: PRT: 593 AA.
AC 064324:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
GN Syntaxin binding protein 2 (UNC-18 homolog 2) (UNC-18B) (MUSEC1).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95197608; PubMed=7890715;
RA Tellam J.T., McIntosh S., James D.E.;
RT "Molecular identification of two novel Munc-18 isoforms expressed in
RT non-neuronal tissues."
RL J. Biol. Chem. 270:5857-5863(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Pancratic Islets;
RA MEDLINE=95197474; PubMed=7890599;
RA Katagiri H., Terasaki J., Murata T., Ishihara H., Ogihara T.,
RA Inukai K., Fukushima Y., Anai M., Kikuchi M., Miyazaki J.I.,
RA Yazaki Y., Oka Y.;
RT "A novel isoform of syntaxin-binding protein homologous to yeast Sec1
RT expressed ubiquitously in mammalian cells."
RL J. Biol. Chem. 270:4963-4966(1995).
CC -I- FUNCTION: INVOLVED IN THE PROTEIN TRAFFICKING FROM THE GOLGI
CC APPARATUS TO THE PLASMA MEMBRANE. BINDS SYNTAXINS 1A AND 4.
CC -I- TISSUE SPECIFICITY: EXPRESSED IN TESTIS, INTESTINE, KIDNEY, RAT

```

```
CC ADIPOSE TISSUE AND 3T3-L1 CELLS.
CC -1- SIMILARITY: BELONGS TO THE STXBP/UNC-18/SECI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D19520; AAA69912.1; -
DR EMBL: D4068; BAA07666.1; -
DR MGD: MGI:107370; Stxbp2.
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1; 1.
KW Protein transport.
SQ SEQUENCE 593 AA; 66357 MW; 1A7735C0566BDC8F CRC64;

Alignment Scores:
Pred. No.: 21.5 Length: 593
Score: 47.00 Matches: 9
Percent Similarity: 77.78% Conservative: 5
Best Local Similarity: 50.00% Mismatches: 4
Query Match: 30.92% Indels: 0
DB: 1 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x STB2_MOUSE (1-593)
OY 20 GACAGTGTGTCGTCTGCTGTCGCTAACCTTAAGACCGCTGTCCTCAG 73
DB 402 ASPLYSLIleAryValleuLeuLeuTyrlleLeuLeuAArgAnslYValSerGlu 419

RESULT 13
STB2_RAT
ID STB2_RAT STANDARD: PRT; 594 AA.
AC Q62753;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE SynTaxin binding protein 2 (Unc-18 homolog 2) (Unc-18B) (Unc18-2).
GN STXB2 OR UNC18B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Liver;
RX MEDLINE=95286582; PubMed=776895;
RA Hata Y., Suedhof T.C.;
RT "A novel ubiquitous form of Munc-18 interacts with multiple
RT syntaxins. Use of the yeast two-hybrid system to study interactions
RT between proteins involved in membrane traffic."
RL J. Biol. Chem. 270:13022-13028(1995).
CC -1- FUNCTION: INVOLVED IN THE PROTEIN TRAFFICKING FROM THE GOLGI
CC APPARATUS TO THE PLASMA MEMBRANE. BINDS SYNNTAXINS 1A, 2, 3 BUT NOT
CC 4.
CC -1- SIMILARITY: BELONGS TO THE STXBP/UNC-18/SECI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U20283; AAA79516.1; -
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1; 1.
KW Protein transport.
SQ SEQUENCE 594 AA; 66695 MW; B4956372A75F1637 CRC64;
```

```
Alignment Scores:
Pred. No.: 21.5 Length: 594
Score: 47.00 Matches: 9
Percent Similarity: 77.78% Conservative: 5
Best Local Similarity: 50.00% Mismatches: 4
Query Match: 30.92% Indels: 0
DB: 1 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x STB2_RAT (1-594)
OY 20 GACAGTGTGTCGTCTGCTGTCGCTAACCTTAAGACCGCTGTCCTCAG 73
DB 402 ASPLYSLIleAryValleuLeuLeuTyrlleLeuLeuAArgAnslYValSerGlu 419

RESULT 14
NU5C_ATRJA
ID NU5C_ATRJA STANDARD: PRT; 749 AA.
AC P51099;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE NADH-plastoquinone oxidoreductase chain 5, chloroplast (EC 1.6.5.3).
GN NDHF.
OS Attractylodes japonica.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Carduaceae;
OC Attractylodes.
OX NCBI_TaxID=41486;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96036088; PubMed=7479788;
RA Kim K.U., Jansen R.K.;
RT "ndhf sequence evolution and the major clades in the sunflower
RT family."
RL Proc. Natl. Acad. Sci. U.S.A. 92:10379-10383(1995).
CC -1- CATALYTIC ACTIVITY: NADH + plastoquinone -> NAD(+) + plastoquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L39413; AAC37727.1; -
DR InterPro: IPR003916; NADHox.reds.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; oxidored_q1; 1.
DR Pfam: PF00662; oxidored_q1_N; 1.
DR Pfam: PF01010; oxidored_q1_C; 1.
DR PRINTS: PR01434; NADHDHGNASE5.
KW Oxidoreductase; NAD; Plastoquinone; Chloroplast.
SQ SEQUENCE 749 AA; 84894 MW; 2C77C7663A2ACA01 CRC64;

Alignment Scores:
Pred. No.: 21.8 Length: 749
Score: 47.00 Matches: 9
Percent Similarity: 57.89% Conservative: 2
Best Local Similarity: 47.37% Mismatches: 8
Query Match: 32.19% Indels: 0
DB: 1 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x NU5C_ATRJA (1-749)
OY 63 CCGCTCTTAAGTGTAGCGACAGACAGACACAACTGCTGTACCTGAGAA 7
DB 525 PROPhelAeuSerIleProHisPheGlyAsnThrIlyeThrIySerIyrrProSerGlu 543
```

RESULT 15
 RPCL_TRYBB STANDARD: PRT: 1530 AA.
 ID RPCL_TRYBB
 AC P08968:
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA-directed RNA polymerase III largest subunit (EC 2.7.7.6).
 OS Trypanosoma brucei brucei.
 CC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 CC NCBI_TaxID=5702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=427 / Isolate MIPat 1.2A;
 RX MEDLINE=89016560; PubMed=3174432;
 RA Cornelissen A.W.C.A., Evers R., Koeck J.;
 RT "Structure and sequence of the gene for the largest subunit of
 RT trypanosomal RNA polymerase III."
 RL Nucleic Acids Res. 16:8753-8772(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90036885; PubMed=2808366;
 RA Smith J.L., Levin J.R., Agabian N.;
 RT "Molecular characterization of the Trypanosoma brucei RNA polymerase
 RT I and III largest subunit genes."
 RL J. Biol. Chem. 264:18091-18099(1989).
 CC -I- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -I- SUBUNIT: RNA POLYMERASE III CONSISTS OF ABOUT 15 DIFFERENT
 CC SUBUNITS. THIS SUBUNIT IS THE LARGEST COMPONENT OF RNA POLYMERASE
 CC III.
 CC -I- SUBCELLULAR LOCATION: Nuclear.
 CC -I- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
 CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
 CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
 CC III FOR 5S AND TRNA GENES.
 CC -I- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 CC DR EMBL: X12494; CAA31014.1; -;
 CC DR EMBL: M27163; AAA30233.1; -;
 CC DR PIR: S01393; S01393.
 CC DR InterPro: IPR000722; RNA_pol_A.
 CC DR InterPro: IPR002879; RNA_pol_A2.
 CC DR Pfam: PF00623; RNA_pol_A; 1.
 CC DR Pfam: PF01854; RNA_pol_A2; 1.
 CC KM Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
 CC zinc-finger; Nuclear protein.
 CC FT ZN_FING 74 87 C3H-TYPE (POTENTIAL).
 CC FT CONFLICT 1325 1325 D -> E (IN REF. 2).
 CC FT CONFLICT 1493 1493 I -> V (IN REF. 2).
 CC SQ SEQUENCE 1530 AA; 170271 MW; FC03D700CEFD742 CRC64;

Alignment Scores:

| | | | |
|------------------------|--------|---------------|------|
| Pred. No.: | 22.9 | Length: | 1530 |
| Score: | 47.00 | Matches: | 7 |
| Percent Similarity: | 66.67% | Conservative: | 3 |
| Best Local Similarity: | 46.67% | Mismatches: | 5 |
| Query Match: | 30.92% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-09-939-293-1_COPY_56_139 (1-84) x RPCL_TRYBB (1-1530)

OY 10 TTCAGGTACAGACAGTGTGTGTGTCCTGTTGTGGCTAACTTT 54
 |||||::: ||| ||| |||:::|
 DB 490 PheatgPheAnsclucCysCysAlaProtyrAsnAlaAspPhe 504

Search completed: February 16, 2003, 03:07:31
 Job time : 10.9925 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: February 16, 2003, 03:05:21 ; Search time 12.2239 seconds
(without alignments)
1321.232 Million cell updates/sec

Title: US-09-939-293-1-COPY_56_139

Perfect score: 152
Sequence: 1 actcattcttcaggtacag.....gtttctcagattgtataga 84

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.SP001/US0939293/r/unat.12022003.170354.8546/app_query.fasta.1.590
-DB=PIR_73 -OFMT=fastan -SUFFIX=pir -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.ccl -LIST=45
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0939293.qcgn.1.1.24-rtunat.12022003.170354.8546 -NCP=6 -ICPU=3
-NO_XLPHY -NO_MMAPP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV_TIMOUT=120
-NARN_TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_73: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 55 | 36.2 | 445 | 2 T31581 | hypothetical prote |
| 2 | 55 | 36.2 | 449 | 2 B85069 | hypothetical prote |
| 3 | 54.5 | 35.9 | 687 | 2 T04927 | probable serine/th |
| 4 | 51 | 34.9 | 268 | 2 AB2622 | type V secretory p |
| 5 | 51 | 34.9 | 2594 | 2 A35774 | kinase-related pro |
| 6 | 50.5 | 34.6 | 4836 | 2 T14346 | herc2 protein - mo |
| 7 | 50 | 32.9 | 432 | 2 A43448 | thrombin receptor |
| 8 | 50 | 34.2 | 744 | 2 T13493 | NADH2 dehydrogenas |
| 9 | 49.5 | 32.6 | 381 | 2 S16506 | hypothetical prote |
| 10 | 49.5 | 33.9 | 621 | 2 T48492 | hypothetical prote |
| 11 | 49.5 | 33.9 | 763 | 2 AC1344 | ribonucleoside-dip |
| 12 | 49.5 | 33.9 | 763 | 2 AG1714 | ribonucleoside-dip |
| 13 | 49 | 32.2 | 366 | 2 I46469 | Mel-1a melatonin r |
| 14 | 49 | 32.2 | 632 | 2 E96723 | hypothetical prote |

| | | | | | |
|----|------|------|------|----------|--------------------|
| 15 | 49 | 32.2 | 637 | 2 T26593 | hypothetical prote |
| 16 | 49 | 33.6 | 744 | 2 T12627 | NADH2 dehydrogenas |
| 17 | 49 | 32.2 | 3924 | 2 S37431 | ankyrin 2, neuroma |
| 18 | 48.5 | 33.2 | 213 | 2 AB2386 | hypothetical prote |
| 19 | 48.5 | 33.2 | 333 | 2 AB2664 | conserved hypothet |
| 20 | 48.5 | 33.2 | 401 | 2 A97446 | hypothetical prote |
| 21 | 48.5 | 33.2 | 773 | 2 I59351 | carnitine O-palmit |
| 22 | 48 | 32.9 | 210 | 2 G84899 | hypothetical prote |
| 23 | 48 | 31.6 | 419 | 2 D82408 | conserved hypothet |
| 24 | 48 | 31.6 | 441 | 2 T50436 | hypothetical prote |
| 25 | 48 | 32.9 | 763 | 1 A49681 | long-chain-fatty-a |
| 26 | 48 | 31.6 | 1691 | 2 D54689 | hypothetical prote |
| 27 | 48 | 31.6 | 1691 | 2 D54689 | protein-tyrosine-p |
| 28 | 48 | 31.6 | 1894 | 2 C54689 | protein-tyrosine-p |
| 29 | 47.5 | 31.2 | 304 | 2 G85068 | N7-like protein [1 |
| 30 | 47.5 | 32.5 | 504 | 2 T21377 | hypothetical prote |
| 31 | 47.5 | 32.5 | 602 | 2 H97563 | glutathione-regula |
| 32 | 47.5 | 32.5 | 602 | 2 AG2784 | hypothetical prote |
| 33 | 47 | 32.2 | 183 | 2 T25711 | hypothetical prote |
| 34 | 47 | 30.9 | 302 | 2 F85068 | N7-like protein [1 |
| 35 | 47 | 30.9 | 583 | 2 A55831 | mu Sec1 protein - |
| 36 | 47 | 30.9 | 594 | 2 A57022 | Munc18-2 - rat |
| 37 | 47 | 32.2 | 741 | 2 T12691 | NADH2 dehydrogenas |
| 38 | 47 | 32.2 | 744 | 2 T13063 | NADH2 dehydrogenas |
| 39 | 47 | 32.2 | 744 | 2 T13755 | NADH2 dehydrogenas |
| 40 | 47 | 32.2 | 749 | 2 T12623 | NADH2 dehydrogenas |
| 41 | 47 | 30.9 | 1388 | 2 T34157 | hypothetical prote |
| 42 | 47 | 30.9 | 1530 | 1 S01393 | DNA-directed RNA p |
| 43 | 47 | 30.9 | 1765 | 2 T42714 | ankyrin 3, splice |
| 44 | 47 | 30.9 | 1912 | 2 A56178 | protein-tyrosine-p |
| 45 | 47 | 30.9 | 1940 | 2 T42715 | ankyrin 3, splice |

ALIGNMENTS

RESULT 1

T31581

hypothetical protein Y105C5A.x - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000

C/Accession: T31581

R/Murray, A.

submitted to the EMBL Data Library, September 1999

A/Reference number: Z21045

A/Accession: T31581

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-445 <MIL>

A/Cross-references: EMBL:AL117193; NID:el549703; PIDN:CA855004.1; CESP:Y105C5A.x

A/Experimental source: clone Y105C5A

C/Genetics:

A/Gene: CESP:Y105C5A.x

A/Introns: 17/3; 71/2; 98/1; 140/2; 267/3; 300/3; 361/3; 400/2

C/Superfamily: basic fibroblast growth factor receptor 1, immunoglobulin homology; pr

Alignment Scores:

Pred. No.: 3.32 Length: 445
Score: 55.00 Matches: 10
Percent Similarity: 64.00% Conservative: 6
Best Local Similarity: 40.00% Mismatches: 0
Query Match: 36.18% Indels: 0
DB: 2 Gaps: 0

US-09-939-293-1-copy_56_139 (1-84) x T31581 (1-445)

QY 1 ACTTCATTTTCAGTACAGACAGTGTTCGTTCCCTTGCGGCTAATTAGAG 60
DB 223 TThArGAsnLeuArGrProHISGIIuIleGluCysAsnProIleuSerAsnPhyTYLys 242
QY 61 CGGTGTTTCACGAA 75
DB 243 ArgCysTyrSerasp 247

F:2222-2499/Domain: protein kinase homology <KIN>
F:2230-2238/Region: protein kinase ATP-binding motif

Alignment Scores:

| | | | |
|------------------------|--------|---------------|------|
| Pred. No.: | 13.6 | Length: | 2594 |
| Score: | 51.00 | Matches: | 9 |
| Percent Similarity: | 69.57% | Conservative: | 7 |
| Best Local Similarity: | 39.13% | Mismatches: | 7 |
| Query Match: | 34.93% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-09-939-293-1_COPY_56_139 (1-84) x A35774 (1-2594)

OY 71 GAGAAACACGCTCTTAAGTTAGCCACAGACACAGACACTGTCTACCTG 12

Db 757 GtndtunsgnleuLeuGtlnleuSerTyrAspIcylYhstlyshsAlaLeuAlaLeu 776

OY 11 AAGAAATGAA 3

Db 777 Serfasnasp 779

RESULT 6

herc2 protein - mouse

N:Alternate names: rjs protein

C:Species: Mus musculus (house mouse)

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 19-May-2000

C:Accession: T14346; T14317

R:Walkiewicz, M.J.; Ji, Y.; Ren, X.; Horsthemke, B.; Francis, F.; Russell, L.B.; Johnson, submitted to the EMBL data library, June 1998

A:Description: Genomic alterations within the Herc2 gene are associated with juvenile le

A:Reference number: Z17996

A:Accession: T14346

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-4836 <EMBL>

A:Cross-references: EMBL:AF071173; NID:94079810; PID:94079811; PIDN:AAD08658.1

Proc. Natl. Acad. Sci. U.S.A. 95, 9436-9441, 1998

A:Title: A very large novel protein with diverse functional motifs is deficient in rjs

A:Reference number: Z17976; MIMD:98356175; PMID:9689098

A:Accession: T14317

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-691; 'L', 693-723, 'A', 725-746, 'R', 748-755, 'L', 757-928, 'R', 930-1113, 'S', 1115-1160, 'T', 3162-3384, 'L', 3387-3507, 'A', 3509-4186, 'S', 4188-4715, 'A', 4717-4722, 'R', 4724-47

A:Cross-references: EMBL:AF061529; NID:93414808; PID:93414809; PIDN:AAC31431.1

A:Experimental source: strain C57BL/6

C:Comment: This protein has several motifs in common with human giant protein p619.

C:Genetics:

A:Gene: Herc2; rjs

A:Map position: 7

C:Superfamily: mouse herc2 protein; ubiquitin-protein ligase homology

Alignment Scores:

| | | | |
|------------------------|--------|---------------|------|
| Pred. No.: | 16 | Length: | 4836 |
| Score: | 50.50 | Matches: | 12 |
| Percent Similarity: | 60.71% | Conservative: | 5 |
| Best Local Similarity: | 42.86% | Mismatches: | 10 |
| Query Match: | 34.59% | Indels: | 1 |
| DB: | 2 | Gaps: | 1 |

US-09-939-293-1_COPY_56_139 (1-84) x T14346 (1-4836)

OY 83 CTATTCATTTCTGGAACACCGCTTCTTA--AAGTTAGCCACACAGACACACAAA 27

Db 4032 leuLeuGtlnleuSerTyrAspIcylYhstlyshsAlaLeuAlaLeu 4051

OY 26 CACTGTCTACTGAAGATGAA 3

Db 4052 HiscysleuAlaLeuSerSerGln 4059

RESULT 7

A43448
thrombin receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: A43448

R:Zhong, C.; Hayzer, D.J.; Corson, M.A.; Runge, M.S.

J. Biol. Chem. 267, 16975-16979, 1992

A:Title: Molecular cloning of the rat vascular smooth muscle thrombin receptor. Evid

A:Reference number: A43448; MIMD:92381002; PMID:1324917

A:Accession: A43448

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-432 <ZHO>

A:Cross-references: GB:M81642; NID:g207465; PIDN:AAA42274.1; PID:g207466

A:Experimental source: RASM aortic smooth muscle cells

A:Note: sequence extracted from NCBI backbone (NCBIN:111973, NCBIP:111974)

C:Keywords: G protein-coupled receptor; transmembrane protein

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-----|
| Pred. No.: | 20.5 | Length: | 432 |
| Score: | 50.00 | Matches: | 11 |
| Percent Similarity: | 55.56% | Conservative: | 4 |
| Best Local Similarity: | 40.74% | Mismatches: | 8 |
| Query Match: | 32.89% | Indels: | 4 |
| DB: | 2 | Gaps: | 1 |

US-09-939-293-1_COPY_56_139 (1-84) x A43448 (1-432)

OY 1 ACTTCATTTCTGACGTACACAGCTGTTGTGCTGTGCTTAAGAAAG 60

Db 355 ThrAlaYrPhaAlaYrLeuGtlnleuSerTyrAspIcylYhstlyshsAlaLeuAlaLeu 371

OY 61 CGGTGTTCTCAGATGATA 81

Db 372 ---CysIleasprProleulle 377

RESULT 8

T13493

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Nassauvia gautichaudii chlor

C:Species: chloroplast Nassauvia gautichaudii

C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 03-Jun-2002

C:Accession: T13493

R:Kim, K.J.; Jansen, R.K.

Proc. Natl. Acad. Sci. U.S.A. 92, 10379-10383, 1995

A:Title: Ndhf sequence evolution and the major clades in the sunflower family.

A:Reference number: Z17549; MIMD:96036088; PMID:7479788

A:Accession: T13493

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-744 <KIM>

A:Cross-references: EMBL:L39405; NID:g845609; PID:g845610; PIDN:AAC37764.1

C:Genetics:

A:Genome: chloroplast

A:Note: ndhf

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5

C:Keywords: chloroplast; membrane-associated complex; NAD; oxidoreductase

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-----|
| Pred. No.: | 20.2 | Length: | 744 |
| Score: | 50.00 | Matches: | 10 |
| Percent Similarity: | 63.16% | Conservative: | 2 |
| Best Local Similarity: | 52.63% | Mismatches: | 7 |
| Query Match: | 34.25% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-09-939-293-1_COPY_56_139 (1-84) x T13493 (1-744)

OY 63 CCGCTTCTTAAGTTAGCCACACAGACACAACTGCTGACCTGAAGAA 7

Db 525 ProPheLeuSerIleSerHisPheGlyAsnThrIlyThrIlySerTyrProTyGln 543

RESULT 9

S16506

OY 20 CTGTACTGAGAAAT 6
||| |||||
Db 555 LeuAlaLeuLysAsn 559

RESULT 13

I46469
Mel-1a melatonin receptor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 17-Mar-2000
C:Accession: I46469
R:Reppert, S.M.; Weaver, D.R.; Ebisawa, T.
Neuron 13, 1177-1185, 1994
A:Title: Cloning and characterization of a mammalian melatonin receptor that mediates re
A:Reference number: 138848; MUID:95033233; PMID:7946354
A:Accession: I46469
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-366 <REP>
A:Cross-references: EMBL:U014109; NID:9602131; PIDN:AA017721.1; PID:9602132
C:Superfamily: vertebrate rhodopsin

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-----|
| Pred. No.: | 29.7 | Length: | 366 |
| Score: | 49.00 | Matches: | 12 |
| Percent Similarity: | 88.24% | Conservative: | 3 |
| Best Local Similarity: | 70.59% | Mismatches: | 0 |
| Query Match: | 32.24% | Indels: | 2 |
| DB: | 2 | Gaps: | 2 |

US-09-939-293-1_COPY_56_139 (1-84) x I46469 (1-366)

OY 11 TCAGTACAGACAGTGTGTG---GTGTCTCTGTGTGG---CTAAGTTTA 55
|||||

Db 156 SerGlyThrAsnSerLeuCysTrpValPheLeuIleTrpThrLeu 172
|||||

RESULT 14

E96723

hypothetical protein F20P5.19 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: E96723

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizart, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, S.X.; Liu, S.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96723

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-632 <SPRO>

A:Cross-references: GB:AE005173; NID:92194130; PIDN:AAB61105.1; GSPDB:GN00141

C:Genetics:

A:Gene: F20P5.19

A:Map position: 1

C:Superfamily: vetispladiene synthase 1

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-----|
| Pred. No.: | 29.3 | Length: | 632 |
| Score: | 49.00 | Matches: | 9 |
| Percent Similarity: | 50.00% | Conservative: | 7 |
| Best Local Similarity: | 28.12% | Mismatches: | 10 |
| Query Match: | 32.24% | Indels: | 6 |
| DB: | 2 | Gaps: | 1 |

US-09-939-293-1_COPY_56_139 (1-84) x E96723 (1-632)

OY 7 TTCTTCAGGTACAGACAGTGTGTGTGTCCTGTGTGCTAACTTAAG-----57

Db 193 PheLeuIlePheHisThrCysMetCysValAspIlePheAsnArgPheLeuGlyAspAsp 212
||| :|||

OY 58 -----AAGCGTGTCTCAGAAATTCATAGA 84
||| :|||

Db 213 GlyAsnPheLeuSerCysLeuAsnAspAspValArg 224
||| :|||

RESULT 15

T26593

hypothetical protein Y32F6A.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26593

R:Barlow, K.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20241

A:Accession: T26593

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-637 <WIL>

A:Cross-references: EMBL:AL021474; PIDN:CA016307.1; GSPDB:GN00023; CESP:Y32F6A.2

A:Experimental source: clone Y32F6A

C:Genetics:

A:Gene: CESP:Y32F6A.2

A:Map position: 5

A:Introns: 45/3; 275/3; 376/2; 453/3; 500/1; 606/3

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-----|
| Pred. No.: | 29.3 | Length: | 637 |
| Score: | 49.00 | Matches: | 12 |
| Percent Similarity: | 59.26% | Conservative: | 4 |
| Best Local Similarity: | 44.44% | Mismatches: | 7 |
| Query Match: | 32.24% | Indels: | 4 |
| DB: | 2 | Gaps: | 1 |

US-09-939-293-1_COPY_56_139 (1-84) x T26593 (1-637)

OY 2 CTTCATCTTACAGTACAGACAGTGTGTGTGTCCTGTGTGCTAAGTAAAGC 61
||| :|||

Db 455 LeuThrAsnSerSerMetGlySerIleTrpValPheLeuPheTrpLeuAlaIa 474
||| :|||

OY 62 -----GGTGTTCCT 70
||| :|||

Db 475 CysSerIleGlnGlyIleSer 481
||| :|||

Search completed: February 16, 2003, 03:11:55
Job time : 15.2239 secs

THIS PAGE BLANK (USPTO)


```
; ORGANISM: Homo sapiens
US-09-796-692-8602

Query Match
Best Local Similarity 100.0%; Score 84; DB 9; Length 549;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTTCATCTTCAGGTACAGACAGAGTGTGTGTCCTGTGTGGCTAACTTTAAGAG 60
    |||
Db 51 ACTTCATCTTCAGGTACAGACAGAGTGTGTGTCCTGTGTGGCTAACTTTAAGAG 110
    |||

OY 61 CGGTGTTTCTCAGAAATTGATAAGA 84
    |||
Db 111 CGGTGTTTCTCAGAAATTGATAAGA 134

RESULT 2
US-09-939-293-1
; Sequence 1, Application US/09939293
; Patent No. US20020132786A1
; GENERAL INFORMATION:
; APPLICANT: Alimonti, Emdad S.
; TITLE OF INVENTION: AN IAP PEPTIDE OR POLYPEPTIDE
; FILE REFERENCE: 480140.465
; CURRENT APPLICATION NUMBER: US/09/939,293
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1358
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)...(739)
US-09-939-293-1

Query Match
Best Local Similarity 100.0%; Score 84; DB 10; Length 1358;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTTCATCTTCAGGTACAGACAGAGTGTGTGTCCTGTGTGGCTAACTTTAAGAG 60
    |||
Db 56 ACTTCATCTTCAGGTACAGACAGAGTGTGTGTCCTGTGTGGCTAACTTTAAGAG 115
    |||

OY 61 CGGTGTTTCTCAGAAATTGATAAGA 84
    |||
Db 116 CGGTGTTTCTCAGAAATTGATAAGA 139

RESULT 3
US-09-925-297-132
; Sequence 132, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antbodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 132
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
```

```
; LOCATION: (63)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (813)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (861)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (885)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-132

Query Match
Best Local Similarity 98.8%; Score 83; DB 10; Length 911;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CTTCATCTTCAGGTACAGACAGAGTGTGTGTCCTGTGTGGCTAACTTTAAGAGC 61
    |||
Db 64 CTTCATCTTCAGGTACAGACAGAGTGTGTGTCCTGTGTGGCTAACTTTAAGAGC 123
    |||

OY 62 GGTGTTTCTCAGAAATTGATAAGA 84
    |||
Db 124 GGTGTTTCTCAGAAATTGATAAGA 146
    |||

RESULT 4
US-09-798-116-3
; Sequence 3, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
; APPLICANT: Verhagen, Anne Marie
; APPLICANT: Ekerl, Paul
; TITLE OF INVENTION: No. US20020110851A1 Polypeptides, Modulatory Agents Therefor
; FILE REFERENCE: 10338-004US
; CURRENT APPLICATION NUMBER: US/09/798,116
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: AU P05995/00
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-798-116-3

Query Match
Best Local Similarity 52.9%; Score 44.4; DB 10; Length 714;
Matches 54; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 15 GTACAGACAGAGTGTGTGTCCTGTGTGGCTAACTTTAAGAGCGGTGTTCCAGA 74
    |||
Db 45 GTTCAGGTACAGACAGCGTGTGTCCTGTGTGGCTAACTTTAAGAGCGGTGTTCCAGA 104
    |||

OY 75 ATTGATTAAGA 84
    |||
Db 105 ATTGATTAAGA 114
    |||

RESULT 5
US-09-798-116-1
; Sequence 1, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
; APPLICANT: Verhagen, Anne Marie
; APPLICANT: Ekerl, Paul
; TITLE OF INVENTION: No. US20020110851A1 Polypeptides, Modulatory Agents Therefor
; FILE REFERENCE: 10338-004US
; CURRENT APPLICATION NUMBER: US/09/798,116
; CURRENT FILING DATE: 2001-03-02
```

```

: PRIOR APPLICATION NUMBER: AU PQ5995/00
:
: PRIOR FILING DATE: 2000-03-02
:
: NUMBER OF SEQ ID NOS: 25
:
: SOFTWARE: PatentIn version 3.0
:
: SEQ ID NO 1
:
:   LENGTH: 1356
:
:   TYPE: DNA
:
:   ORGANISM: Mus musculus
:
:   FEATURE:
:
:     NAME/KEY: 5'UTR
:
:     LOCATION: (1)..(10)
:
:     NAME/KEY: 3'UTR
:
:     LOCATION: (725)..(1356)
:
: OS-09-798-116-1

```

| | | | | |
|--------------------------|--------|--------------------|-----------|--------------|
| Query Match | 52.9% | Score 44.4; | DB 10; | Length 1356; |
| Best Local Similarity | 77.1%; | Pred. No. 8.7e-06; | | |
| Matches 54; Conservative | 0; | Mismatches 16; | Indels 0; | Gaps 0; |

OY 15 GTACAGACAGTGTTTGTTGTTCCCTGTTTGCGCAACTTTTAAGAAGCGGTGTTTTCTAGA 74
||| | | ||||||| ||||| ||||| | |||||||
Db 55 GTTCAGGTACAGACAGCGTTTCCCTGCTTGCGCAACTCTAAGAAGCGGTGTTTTCTAGA 110

| | | | |
|----|-----|------------|-----|
| QY | 75 | ATTGATAAGA | 84 |
| | | | |
| Db | 115 | ATTGATAAAA | 124 |

RESULT 6
US-09-800-631-96

```

? Sequence 96/Application US/09800631
? Patent No. US20020082228A1
? GENERAL INFORMATION:
? APPLICANT: Hong Zhang
? TITLE OF INVENTION: ANTISENSE MODULATION OF BH3 INTERACTING DOMAIN DEATH AGONIST EXPRESSION
? FILE REFERENCE: ISPH-0544
? CURRENT APPLICATION NUMBER: US/09/800,631
? CURRENT FILING DATE: 2001-03-07
? PRIOR APPLICATION NUMBER: US/09/657,346
? PRIOR FILING DATE: 2000-09-07
? NUMBER OF SEQ ID NOS: 175
? SEQ ID NO 96
? LENGTH: 30310
? TYPE: DNA
? ORGANISM: Mus musculus
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (19791)...(19802)
? NAME/KEY: CDS
? LOCATION: (21160)...(21370)
? NAME/KEY: CDS
? LOCATION: (24168)...(24307)
? NAME/KEY: CDS
? LOCATION: (25696)...(25908)
? NAME/KEY: CDS
? LOCATION: (27235)...(27246)
? OS-09-800-631-96

```

| | | | | |
|--------------------------|--------|----------------|-----------|---------------|
| Query Match | 30.5%; | Score 25.6; | DB 10; | Length 30310; |
| Best Local Similarity | 59.7%; | Pred. No. 38; | | |
| Matches 43; Conservative | 0; | Mismatches 29; | Indels 0; | Gaps 0 |

0y 7 TTCTTACAGGTACAGACAGTGTTGTCGTCTCGTTGTGGCTAACTTTAAGAAGCGGTG 66
||| | ||| | ||| | ||| | ||| | ||| |
Db 18413 TTCGCATTACTGAACACTTGTGGGGGTTGGGCCGGCCAGGTTTCAAGACAGGTTT 18472

| | | | |
|----|-------|--------------|-------|
| QY | 67 | TTCTCAGATTG | 78 |
| | | | |
| Db | 18473 | TTCTCTGAATAG | 18484 |

RESULT 7

```

US-09-867-701-3333/c
: Sequence 3333, Application US/09867701
: Patent No. US2002012237A1
: GENERAL INFORMATION:
: APPLICANT: Aglate, Paul A.
: APPLICANT: Jones, Robert
: APPLICANT: Harlocker, Susan L.
: TITLE OF INVENTION: COMPOSITIONS AND METH
: TITLE OF INVENTION: AND DIAGNOSIS OF OVA
: FILE REFERENCE: 210121.497
: CURRENT APPLICATION NUMBER: US/09/867,701
: CURRENT FILING DATE: 2001-05-29
: NUMBER OF SEQ ID NOS: 10912
: SOFTWARE: FastSeq for Windows Version 4.0.
: SEQ ID NO 3333
: LENGTH: 223
: TYPE: DNA
: ORGANISM: Homo sapien
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(223)
: OTHER INFORMATION: n = A,T,C or G
US-09-867-701-3333

```

| | | | | |
|-----------------------|-----------------|----------------|-----------|-------------|
| Query Match | 30.0%; | Score 25.2; | DB 10; | Length 223; |
| Best Local Similarity | 60.0%; | Pred. No. 14; | | |
| Matches 42; | Conservative 0; | Mismatches 28; | Indels 0; | Gaps 0; |

Qy 4 TCATCTTCAGGTACAGACAGACTTTGCTGTCCTGTGGCTAACCTTAAAGAACC GG 63
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 TCATACTTAGTGACGCCAGGTTTGCTTCAAGTGATACCATCGGTAAAAATCCTC 140

| | | | |
|----|-----|------------|-----|
| QY | 64 | TGTTTCTCAG | 73 |
| | | | |
| Db | 139 | TTTTCTCTCG | 130 |

```

RESULT 8
US-09-764-864-367/c
: Sequence 367. Application US/09764864
: Patent No. US20020132753A1
:
GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P123
: CURRENT APPLICATION NUMBER: US/09/764,864
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1792
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 367
:
LENGTH: 1695
:
TYPE: DNA
:
ORGANISM: Homo sapiens
US-09-764-864-367

```

| | | | | |
|-----------------------|-----------------|----------------|-----------|--------------|
| Query Match | 30.0%; | Score 25.2; | DB 10; | Length 1695; |
| Best Local Similarity | 60.0%; | Pred. No. 24; | | |
| Matches 42; | Conservative 0; | Mismatches 28; | Indels 0; | Gaps 0; |

```

Oy      4 TCATCTTTCAGGTTACAGACAGTGTTGTGTGTCCTGTTGTGGCTTAACCTTTAAGAAGCCG 63
        ||||| ||| ||| ||| ||||| || ||||| || | ||| |||
Db     386 TCATACTTAGTGTCAGGCCAGCGGTTGTCTTCACAAGTTGATATACCATCGGTAATAATGCCTC 32

```

| | | | |
|----|-----|------------|-----|
| QY | 64 | TGTTTCTCAG | 73 |
| | | | |
| Db | 326 | TTTTCTTCTG | 317 |

RESULT 9
US-09-764-847-1190
; Sequence 1190, Application US/09764847
; Patent No. US20020132767A1

```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1190
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1190

Query Match      29.8%; Score 25; DB 10; Length 536;
Best Local Similarity 61.5%; Pred. No. 21;
Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY      11 TCAGGTACAGACAGAGTGTGTGTCCTCGTGTGGCTACTTAAAGACCGGTGTTCT 70
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      438 TCACGTCCAGAGATTATTCTCTCTCCCTGCTACAGACCCTGTAAAGTTTGGAGCTTT 487

OY      71 CAGAA 75
         |||||
Db      488 CAGAA 492

RESULT 10
US-09-764-864-780/c
; Sequence 780, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 780
; LENGTH: 905
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (168)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (190)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (775)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (800)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (836)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (849)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (875)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (892)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (900)
; OTHER INFORMATION: n equals a,t,g, or c

```

```

US-09-764-864-780
Query Match
Best Local Similarity 29.5%; Score 24.8; DB 10; Length 905;
Matches 41; Conservative 1; Mismatches 28; Indels 0; Gaps 0;

QY 4 TCATTCTTCAGGTACAGACAGCTGTGTGTCCTGTTGGCTAACTTTAAGAACGG 63
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 391 WCATCTTAAAGTGTAGACGGCCAGGGTTGTCCTCAAGTTGATACCATCGGTAAATGCTC 332

QY 64 TGTTCCTCAG 73
      | || |||
DB 331 TTTCTCTCTG 322

RESULT 11
US-09-731-231A-3
; Sequence 3, Application US/09731231A
; Patent No. US20020082189A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001007
; CURRENT APPLICATION NUMBER: US/09/731,231A
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 326014
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(326014)
; OTHER INFORMATION: n - A,T,C or G
US-09-731-231A-3

Query Match
Best Local Similarity 29.5%; Score 24.8; DB 10; Length 326014;
Matches 44; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 4 TCATTCTTCAGGTACAGACAGCTGTGTGTCCTGTTGGCTAACTTTAAGAACGG 63
      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 238934 TCATTCTTCAGGTAAAGTAAAGATTAATTTATTTGATTTGCCAAGTTTAAAGACGCT 238993

QY 64 TGTTCCTCAGATGA 79
      || || ||| ||| |||
DB 238994 TGCATCACAGGAATGA 239009

RESULT 12
US-09-938-842A-2118/C
; Sequence 2118, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kireps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379

```


| | | | |
|----|-------|---|-------|
| QY | 4 | TCAATCTCAGGACAGAGAGTGTGTGTCCTGTGTGGGTAACTTTAAGAACGG | 63 |
| | | | |
| | | | |
| | | | |
| | | | |
| Db | 26683 | TCCTTTCTTCAGGACACACCATCTGCTTTTCCCTGTGGAGAAACATTTGCACACT | 26622 |
| | | | |
| | | | |
| | | | |
| | | | |
| QY | 64 | TGTTTCTCAGAAATTGATAA | 82 |

Wed Feb 19 10:49:04 2003

us-09-939-293-1_copy_56_139.rnpb

Page 6

Db 26623 TCTGAACAAATTTCAGAA 26605

Search completed: February 16, 2003, 03:04:48
Job time : 139.433 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: February 16, 2003, 03:12:06 ; Search time 6.89552 Seconds

(without alignments)
622.464 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 152

Sequence: 1 actctccttcagctacgag.....gttcctcagatgtataga 84

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 140259 segs, 25548876 residues

Total number of hits satisfying chosen parameters: 280518

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+np.model -DEV=rlp

-Q=/cg2_1/USPTO.spool/US09393293/runat_12022003_170356.8661/app_query.fasta_1.590

-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1

-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS-human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -USER=US09393293_86CGN_1_1_6_@runat_12022003_170356.8661

-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MAMP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_AA:*

1: /cg2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

2: /cg2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cg2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cg2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cg2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cg2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

7: /cg2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

8: /cg2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cg2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*

10: /cg2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

11: /cg2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*

12: /cg2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*

13: /cg2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

14: /cg2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-------------------|---------------------|
| 1 | 147 | 96.7 | 227 | 10 | US-09-925-297-591 | Sequence 591, Appli |
| 2 | 87 | 57.2 | 237 | 10 | US-09-798-116-2 | Sequence 2, Appli |
| 3 | 87 | 57.2 | 237 | 10 | US-09-798-116-4 | Sequence 4, Appli |
| 4 | 79 | 52.0 | 84 | 10 | US-09-798-116-9 | Sequence 9, Appli |

| | | | | | | |
|----|------|------|------|----|---------------------|--------------------|
| 5 | 49 | 32.2 | 366 | 9 | US-09-992-331-8 | Sequence 8, Appli |
| 6 | 49 | 32.2 | 366 | 9 | US-09-992-331-9 | Sequence 9, Appli |
| 7 | 48.5 | 31.9 | 155 | 9 | US-09-981-876-164 | Sequence 164, App |
| 8 | 48.5 | 31.9 | 155 | 9 | US-09-981-876-225 | Sequence 225, App |
| 9 | 48.5 | 31.9 | 155 | 9 | US-09-148-545-164 | Sequence 164, App |
| 10 | 48.5 | 31.9 | 155 | 9 | US-09-148-545-225 | Sequence 225, App |
| 11 | 48 | 32.9 | 191 | 10 | US-09-811-284-257 | Sequence 257, App |
| 12 | 47 | 30.9 | 166 | 10 | US-09-817-910-9 | Sequence 9, Appli |
| 13 | 47 | 30.9 | 171 | 10 | US-09-798-028-28 | Sequence 28, Appli |
| 14 | 46 | 30.3 | 76 | 9 | US-09-796-692-1217 | Sequence 1217, Ap |
| 15 | 46 | 30.3 | 76 | 9 | US-09-796-692-1676 | Sequence 1676, Ap |
| 16 | 45 | 30.3 | 76 | 9 | US-09-796-692-2235 | Sequence 2235, Ap |
| 17 | 45 | 30.8 | 52 | 9 | US-10-001-887-91 | Sequence 91, Appli |
| 18 | 45 | 30.8 | 131 | 10 | US-09-881-752A-326 | Sequence 326, App |
| 19 | 45 | 30.8 | 377 | 9 | US-09-738-626-4823 | Sequence 4823, Ap |
| 20 | 45 | 29.6 | 389 | 9 | US-09-738-626-4492 | Sequence 4492, Ap |
| 21 | 45 | 30.5 | 452 | 10 | US-09-925-297-663 | Sequence 683, App |
| 22 | 44.5 | 30.5 | 200 | 10 | US-09-912-020-359 | Sequence 359, App |
| 23 | 44.5 | 30.5 | 721 | 10 | US-09-727-801-14 | Sequence 14, Appli |
| 24 | 44 | 28.9 | 66 | 10 | US-09-925-298-873 | Sequence 873, App |
| 25 | 44 | 28.9 | 108 | 10 | US-09-814-122-72 | Sequence 72, Appli |
| 26 | 44 | 28.9 | 245 | 10 | US-09-796-856-20 | Sequence 20, Appli |
| 27 | 44 | 30.1 | 759 | 9 | US-09-975-719-170 | Sequence 170, App |
| 28 | 44 | 30.1 | 759 | 10 | US-09-815-242-5060 | Sequence 5060, Ap |
| 29 | 44 | 30.1 | 2835 | 10 | US-09-885-535-4 | Sequence 4, Appli |
| 30 | 44 | 28.9 | 3075 | 10 | US-09-938-275-5 | Sequence 5, Appli |
| 31 | 43.5 | 29.8 | 85 | 10 | US-09-867-550-878 | Sequence 878, App |
| 32 | 43.5 | 29.8 | 116 | 9 | US-09-895-913A-178 | Sequence 178, App |
| 33 | 43.5 | 29.8 | 116 | 10 | US-09-815-242-11582 | Sequence 11582, A |
| 34 | 43 | 29.5 | 64 | 9 | US-09-854-133-74 | Sequence 74, Appli |
| 35 | 43 | 29.5 | 64 | 10 | US-09-738-973-74 | Sequence 74, Appli |
| 36 | 43 | 29.5 | 97 | 10 | US-09-864-761-40638 | Sequence 40638, A |
| 37 | 43 | 29.5 | 138 | 10 | US-09-864-761-37139 | Sequence 37139, A |
| 38 | 43 | 29.5 | 143 | 10 | US-09-764-877-1657 | Sequence 1657, Ap |
| 39 | 43 | 28.3 | 211 | 9 | US-09-738-626-4604 | Sequence 4604, Ap |
| 40 | 43 | 28.3 | 213 | 10 | US-09-796-858-22 | Sequence 22, Appli |
| 41 | 43 | 29.5 | 326 | 10 | US-09-953-956-2 | Sequence 2, Appli |
| 42 | 43 | 29.5 | 329 | 10 | US-09-953-956-2 | Sequence 2, Appli |
| 43 | 43 | 29.5 | 329 | 12 | US-10-114-464-2 | Sequence 2, Appli |
| 44 | 43 | 29.5 | 361 | 10 | US-09-925-300-1693 | Sequence 1693, Ap |
| 45 | 43 | 29.5 | 497 | 9 | US-09-866-570A-65 | Sequence 65, Appli |

ALIGNMENTS

RESULT 1
US-09-925-297-591
; Sequence 591, Application US/09925297
; Patent No. US20020081659A1
GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL05
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 591
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-591
Alignment Scores: 8.71e-15 Length: 227

Score: 147.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.71% Indels: 0
DB: 10 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x US-09-925-297-591 (1-227)

OY 4 TCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAGCGG 63
||| |||||||||||||||||||
DB 2 SerPhePheArgTyrArgGlnCysValProValValAlaAsnPhelysLysArg 21

OY 64 TGTTCCTCAGAAATTGATTAAGA 84
|||||
DB 22 CysPheSerGlnLeuIleLys 28

RESULT 2

US-09-798-116-2
; Sequence 2, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
; APPLICANT: Verhagen, Anne Marie
; APPLICANT: Ekert, Paul
; APPLICANT: Vaux, David

; TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor an
; FILE REFERENCE: 10338-004US
; CURRENT APPLICATION NUMBER: US/09/798,116
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: AU P05995/00
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-798-116-2

Alignment Scores:
Pred. No.: 1.51e-05
Score: 87.00
Percent Similarity: 77.78%
Best Local Similarity: 70.37%
Query Match: 57.24%
DB: 10
Matches: 237
Conservative: 19
Mismatches: 2
Indels: 4
Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x US-09-798-116-2 (1-237)

OY 4 TCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAGCGG 63
||| |||||||||||||||||||
DB 14 SerLeuPheArgTyrArgGlnArgPhe-----ProValLeuAlaAsnSerLysLysArg 31

OY 64 TGTTCCTCAGAAATTGATTAAGA 84
|||||
DB 32 CysPheSerGlnLeuIleLys 38

RESULT 3

US-09-798-116-4
; Sequence 4, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
; APPLICANT: Verhagen, Anne Marie
; APPLICANT: Ekert, Paul
; APPLICANT: Vaux, David

; TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor an
; FILE REFERENCE: 10338-004US
; CURRENT APPLICATION NUMBER: US/09/798,116
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: AU P05995/00
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4

; LENGTH: 237
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-798-116-4

Alignment Scores:

Pred. No.: 1.51e-05 Length: 237
Score: 87.00 Matches: 19
Percent Similarity: 77.78% Conservative: 2
Best Local Similarity: 70.37% Mismatches: 4
Query Match: 57.24% Indels: 2
DB: 10 Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x US-09-798-116-4 (1-237)

OY 4 TCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAGCGG 63
||| |||||||||||||||||||
DB 14 SerLeuPheArgTyrArgGlnArgPhe-----ProValLeuAlaAsnSerLysLysArg 31

OY 64 TGTTCCTCAGAAATTGATTAAGA 84
|||||
DB 32 CysPheSerGlnLeuIleLys 38

RESULT 4

US-09-798-116-9
; Sequence 9, Application US/09798116
; Patent No. US20020110851A1
; GENERAL INFORMATION:
; APPLICANT: Verhagen, Anne Marie
; APPLICANT: Ekert, Paul
; APPLICANT: Vaux, David

; TITLE OF INVENTION: No. US20020110851A1el Polypeptides, Modulatory Agents Therefor
; FILE REFERENCE: 10338-004US
; CURRENT APPLICATION NUMBER: US/09/798,116
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: AU P05995/00
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-798-116-9

Alignment Scores:

Pred. No.: 0.000228 Length: 84
Score: 79.00 Matches: 19
Percent Similarity: 67.86% Conservative: 0
Best Local Similarity: 67.86% Mismatches: 7
Query Match: 51.97% Indels: 2
DB: 10 Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x US-09-798-116-9 (1-84)

OY 1 ACTTCATTCTTCAGGTACAGACAGTGTGTGTCCTGTTGGCTTAACCTTAAGAAG 60
||| ||||||| ||| |||||||
DB 13 ThrPheLeuPheArgTyrArgGlnArgPhe-----ProValSerAlaAsnSerLysLys 30

OY 61 CGGTGTTTCAGAAATTGATTAAGA 84
|||||
DB 31 ArgCysPheSerGlnLeuIleLys 38

RESULT 5

US-09-992-331-8
; Sequence 8, Application US/09992331
; Publication No. US20030022186A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, JOHN N.
; APPLICANT: MINTIER, GABE
; APPLICANT: RAMANATHAN, CHANDRA S.
; APPLICANT: HAWKEN, DONALD R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBM18,

```

; TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
; FILE REFERENCE: D0048NP
; CURRENT APPLICATION NUMBER: US/09/992,331
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/308,540
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/261,782
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/248,483
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Ovis aries
US-09-992-331-8

Alignment Scores:
Pred. No.: 11.2 Length: 366
Score: 49.00 Matches: 12
Percent Similarity: 88.24% Conservative: 3
Best Local Similarity: 70.59% Mismatches: 0
Query Match: 32.24% Indels: 2
DB: 9 Gaps: 2

US-09-939-293-1_COPY_56_139 (1-84) x US-09-992-331-8 (1-366)
QY 11 TCAGGTACAGACAGTGTGTGT---GNGTTCCTGTGTGTG---CTAACTTAA 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 156 SerGlyThrnsSerLeucCysTyrValPheLeuIleTrpHrLeuThrIleu 172

RESULT 6
US-09-992-331-9
; Sequence 9, Application US/09992331
; Publication No. US20030022186A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, JOHN N.
; APPLICANT: MINTER, GABE
; APPLICANT: RAMANATHAN, CHANDRA S.
; APPLICANT: HAMKEN, DONALD R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGRPMY18,
; TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: D0048NP
; CURRENT APPLICATION NUMBER: US/09/992,331
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/308,540
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/261,782
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/248,483
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Ovis aries
US-09-992-331-9

Alignment Scores:
Pred. No.: 11.2 Length: 366
Score: 49.00 Matches: 12
Percent Similarity: 88.24% Conservative: 3
Best Local Similarity: 70.59% Mismatches: 0
Query Match: 32.24% Indels: 2
DB: 9 Gaps: 2

US-09-939-293-1_COPY_56_139 (1-84) x US-09-992-331-9 (1-366)
QY 11 TCAGGTACAGACAGTGTGTGT---GNGTTCCTGTGTGTG---CTAACTTAA 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 156 SerGlyThrnsSerLeucCysTyrValPheLeuIleTrpHrLeuThrIleu 172
```

```

Db 156 SerGlyThrnsSerLeucCysTyrValPheLeuIleTrpHrLeuThrIleu 172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 7
US-09-981-876-164
; Sequence 164, Application US/09981876
; Patent No. US20020164669A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/981,876
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,161
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,581
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,584
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,500
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,613
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,582
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,596
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,612
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,632
; PRIOR FILING DATE: 1997-05-23
```


RESULT 8
US-09-981-876-225
; Sequence 225, Application US/09981876
; Patent No. US2002016469A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001PI
; CURRENT APPLICATION NUMBER: US/09/981,876
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,161
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,581
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,584
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,500
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,613
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,582
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,596
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,612
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,601
; PRIOR FILING DATE: 1997-05-23

; PRIOR APPLICATION NUMBER: 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,568
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,569
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,311
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,671
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,674
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,669
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,312
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,313
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,672
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,315
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/056,886
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,877
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,889
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,893
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,630
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,878
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,662
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,872
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,882
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,637
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,903
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,888
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,879
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,880
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,894
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,911
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,636
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,874
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,910
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,864
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,631
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,892
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/047,595

```
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: 60/057,761
/ PRIOR FILING DATE: 05-Sep-1997
/ PRIOR APPLICATION NUMBER: 60/047,599
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: 60/047,588
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: 60/047,585
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: 60/047,586
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: 60/047,590
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: 60/047,594
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: 60/047,589
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: 60/043,578
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: 60/043,576
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: 60/047,501
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: 60/043,670
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: 60/056,632
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 60/056,664
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 60/056,876
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 60/056,881
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 60/056,909
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 60/056,875
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 60/056,862
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 60/056,887
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 60/056,908
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: 60/048,964
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 60/057,650
/ PRIOR FILING DATE: 1997-09-05
/ PRIOR APPLICATION NUMBER: 60/056,884
/ PRIOR FILING DATE: 1997-08-22
/ NUMBER OF SEQ ID NOS: 280
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 225
/ LENGTH: 155

Alignment Scores:
Pred. No.: 12.1 Length: 155
Score: 48.50 Matches: 9
Percent Similarity: 64.71% Conservative: 2
Best Local Similarity: 52.94% Mismatches: 3
Query Match: 31.91% Indels: 3
DB: 9 Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x US-09-981-876-225 (1-155)
OY 3 TTGATCTCAGGTACAGACAGTGTGTTGTTGTCGCTGTTGT 44
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 14 phevalluservai|a|thritiellie|ecysp|h|cys|ser|cys|cs 30
RESULT 9
```

```
US-09-148-545-164
/ Sequence 164, Application US/09148545
/ Publication No. US20030027132A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 70 Human Secreted Proteins
/ FILE REFERENCE: P2001P1
/ CURRENT APPLICATION NUMBER: US/09/148,545
/ CURRENT FILING DATE: 1998-09-04
/ EARLIER APPLICATION NUMBER: PCT/US98/04482
/ EARLIER FILING DATE: 1998-03-06
/ EARLIER APPLICATION NUMBER: 60/040,162
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,333
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/038,621
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,161
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,626
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,334
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,336
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,163
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/047,615
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,600
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,597
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,502
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,633
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,583
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,617
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,618
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,503
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,592
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,581
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,584
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,500
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,587
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,492
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,598
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,613
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,582
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,596
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,612
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,632
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,601
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/043,580
/ EARLIER FILING DATE: 1997-04-11
```



```
Publication No. US20030027132A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 70 Human Secreted Proteins
FILE REFERENCE: P2001P1
CURRENT APPLICATION NUMBER: US/09/148,545
EARLIER FILING DATE: 1998-09-04
EARLIER APPLICATION NUMBER: PCT/US98/04482
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,161
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,517
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 05-Sep-1997
EARLIER APPLICATION NUMBER: 60/047,599
```

```
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 225
; LENGTH: 155

Alignment Scores:
Pred. No.: 12.1 Length: 155
Score: 48.50 Matches: 9
Percent Similarity: 64.71% Conservative: 2
Best Local Similarity: 52.94% Mismatches: 3
Query Match: 31.91% Indels: 3
DB: Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x US-09-148-545-225 (1-155)

Oy 3 TTCATCTTCAGGTACAGACAGTG-----TTTGTCGTTCCTGTGT 44
    |||:||||| ||| |||:| ||| ||| ||| ||| |||
Db 14 PheValIeuSerValIValThrIleIleIleCysPheThrCysSerCys 30

RESULT 11
US-09-811-284-257
; Sequence 257, Application US/09811284
; Patent No. US2002005306A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Vogeli, Gabriel
; TITLE OF INVENTION: No. US2002005306A1e1 G Protein-Coupled Receptors
; FILE REFERENCE: 00167US1
; CURRENT APPLICATION NUMBER: US/09/811,284
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,783
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,907
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,918
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,960
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,917
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/192,945
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,916
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,923
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,933
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,830
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,234
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,155
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,935
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 257
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-811-284-257

Alignment Scores:
Pred. No.: 14.8 Length: 191
Score: 48.00 Matches: 8
Percent Similarity: 72.22% Conservative: 5
Best Local Similarity: 44.44% Mismatches: 5
Query Match: 32.88% Indels: 0
DB: Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x US-09-811-284-257 (1-191)

Oy 68 AACACCGCCTCTTAAAGTTAGCCACAGACACACACAAACTGTCTGTAC 15
    ::||| ||| ||| |||: ||| ||| ||| ||| |||
Db 55 ArgHisArgPheLeuGlnAsnSerIleuThrIysSerHisIysMetCysArgPhe 72

RESULT 12
US-09-817-910-9
; Sequence 9, Application US/09817910
; Patent No. US20020016449A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: 46743 and 27417 NOVEL HUMAN
; FILE REFERENCE: 10448-032001
; CURRENT APPLICATION NUMBER: US/09/817,910
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
```

```
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-817-910-9

Alignment Scores:
Pred. No.: 20.8      Length: 166
Score: 47.00        Matches: 11
Percent Similarity: 54.17%  Conservative: 2
Best Local Similarity: 45.83%  Mismatches: 9
Query Match: 30.92%      Indels: 2
DB: 10              Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x US-09-817-910-9 (1-166)
QY 5 CATCTTCAGGTACAGACAGTGTGTGTTCTGTTGGCTAACTTTAAGAAGCGGT 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 29 HisGlnserIyrlEuasp-----TrrIlePheLeuIrrPrrLeuAlaIyrrArGserGly 46
QY 65 GTTTCACGAAT 76
   ::| |||
Db 47 LeuGlyAlaasn 50

RESULT 13
US-09-798-029-28
; Sequence 28, Application US/09798029
; Patent No. US20020019030A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; APPLICANT: Macdolph, Kyle J.
; APPLICANT: Rudolph-Owen, Laura
; TITLE OF INVENTION: 25324, 50287, 28899, 47007, and 42967
; TITLE OF INVENTION: TRANSFERASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 38155-20001.00
; CURRENT APPLICATION NUMBER: US/09/798,029
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/185,711
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid
US-09-798-029-28

Alignment Scores:
Pred. No.: 20.9      Length: 171
Score: 47.00        Matches: 11
Percent Similarity: 54.17%  Conservative: 2
Best Local Similarity: 45.83%  Mismatches: 9
Query Match: 30.92%      Indels: 2
DB: 10              Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x US-09-798-029-28 (1-171)
QY 5 CATCTTCAGGTACAGACAGTGTGTGTTCTGTTGGCTAACTTTAAGAAGCGGT 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35 HisGlnserIyrlEuasp-----TrrIlePheLeuIrrPrrLeuAlaIyrrArGserGly 52
QY 65 GTTTCACGAAT 76
   ::| |||
Db 53 LeuGlyAlaasn 56

RESULT 14
US-09-796-692-1217
; Sequence 1217, Application US/09796692
; Publication No. US20020198362A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THE
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077,001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1217
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1217

Alignment Scores:
Pred. No.: 27.1      Length: 76
Score: 46.00        Matches: 8
Percent Similarity: 76.92%  Conservative: 3
Best Local Similarity: 61.54%  Mismatches: 2
Query Match: 30.26%      Indels: 0
DB: 9              Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x US-09-796-692-1217 (1-76)
QY 16 TACAGACAGTGTGTGTTCTGTTGGCTAACTTT 54
   ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 TrpHisIyScyStyrcysValGlyValAlaIaIaaspHe 45

RESULT 15
US-09-796-692-1676
; Sequence 1676, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THE
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077,001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1217
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1217
```

```
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1676
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-1676
```

```
Alignment Scores:
Pred. No.:      27.1      Length:      76
Score:          46.00     Matches:      8
Percent Similarity: 76.92% Conservative: 2
Best Local Similarity: 61.54% Mismatches: 3
Query Match:     30.26%   Indels:      0
DB:              9       Gaps:       0
```

US-09-939-293-1_COPY_56_139 (1-84) x US-09-796-692-1676 (1-76)

```
OY      16  TACAGACAGCTGTGTGTCCTGTCGTGCGCTAAGTTT 54
      ::  ::::| | | | | | | | | | | | | | | |
Db      33  TTPHSLYSCYSTYRCYSVALGlyVALVALAAsnpHe 45
```

Search completed: February 16, 2003, 03:28:17
Job time : 7.89552 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 16, 2003, 03:06:36 ; Search time 9.71642 Seconds
(without alignments)
508.732 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 152
Sequence: 1 actctctcttcaggtacag.....gtttctcagaattgataaga 84

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=x1p
-Q=/cgn2_1/USPTO.spool/US0939293/runat_12022003_170354_8576/app_query.fasta_1.590
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human0.cdl
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US0939293.qcgn2_1_17_@runat_12022003_170354_8576 -NCPD=6 -ICPD=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCOTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 152 | 100.0 | 239 | 3 | US-09-479-309-2 |
| 2 | 62 | 40.8 | 190 | 4 | US-09-134-001C-3138 |
| 3 | 49 | 32.2 | 69 | 4 | US-09-082-059-3 |
| 4 | 49 | 32.2 | 366 | 2 | US-08-466-103A-4 |
| 5 | 49 | 32.2 | 366 | 2 | US-08-896-365-7 |
| 6 | 49 | 32.2 | 1839 | 2 | US-09-172-977-4 |
| 7 | 48.5 | 31.9 | 137 | 4 | US-09-247-155-109 |
| 8 | 47 | 32.2 | 267 | 4 | US-08-558-935-2 |
| 9 | 47 | 30.9 | 593 | 2 | US-08-900-927-1 |
| 10 | 47 | 30.9 | 593 | 2 | US-08-900-927-3 |
| 11 | 47 | 30.9 | 593 | 2 | US-08-900-927-4 |
| 12 | 47 | 30.9 | 593 | 2 | US-09-191-279-1 |

| | | | | | | |
|----|------|------|------|---|---------------------|--------------------|
| 13 | 47 | 30.9 | 593 | 2 | US-09-191-279-3 | Sequence 3, Appl1 |
| 14 | 47 | 30.9 | 593 | 2 | US-09-191-279-4 | Sequence 4, Appl1 |
| 15 | 47 | 30.9 | 593 | 4 | US-09-334-476-1 | Sequence 1, Appl1 |
| 16 | 47 | 30.9 | 593 | 4 | US-09-334-476-3 | Sequence 3, Appl1 |
| 17 | 47 | 30.9 | 593 | 4 | US-09-334-476-4 | Sequence 4, Appl1 |
| 18 | 47 | 30.9 | 1088 | 4 | US-09-082-059-2 | Sequence 2, Appl1 |
| 19 | 46 | 30.3 | 69 | 4 | US-09-082-059-7 | Sequence 7, Appl1 |
| 20 | 46 | 30.3 | 69 | 4 | US-09-082-059-13 | Sequence 13, Appl1 |
| 21 | 46 | 30.3 | 522 | 3 | US-09-165-042-3 | Sequence 3, Appl1 |
| 22 | 46 | 31.5 | 606 | 4 | US-09-292-858B-20 | Sequence 20, Appl1 |
| 23 | 45.5 | 29.9 | 1732 | 2 | US-08-477-451-14 | Sequence 14, Appl1 |
| 24 | 45 | 29.6 | 69 | 4 | US-09-082-059-9 | Sequence 9, Appl1 |
| 25 | 45 | 30.8 | 372 | 2 | US-08-837-593-9 | Sequence 9, Appl1 |
| 26 | 44 | 28.9 | 25 | 4 | US-09-336-536-52 | Sequence 52, Appl1 |
| 27 | 44 | 28.9 | 56 | 4 | US-09-336-536-55 | Sequence 55, Appl1 |
| 28 | 44 | 28.9 | 69 | 4 | US-09-082-059-10 | Sequence 10, Appl1 |
| 29 | 44 | 28.9 | 69 | 4 | US-09-082-059-11 | Sequence 11, Appl1 |
| 30 | 44 | 28.9 | 69 | 4 | US-09-082-059-12 | Sequence 12, Appl1 |
| 31 | 44 | 28.9 | 69 | 4 | US-09-082-059-14 | Sequence 14, Appl1 |
| 32 | 44 | 28.9 | 112 | 1 | US-08-477-877B-87 | Sequence 87, Appl1 |
| 33 | 44 | 28.9 | 112 | 2 | US-08-472-281A-87 | Sequence 87, Appl1 |
| 34 | 44 | 28.9 | 112 | 2 | US-08-477-898B-87 | Sequence 87, Appl1 |
| 35 | 44 | 28.9 | 132 | 1 | US-08-477-877B-84 | Sequence 84, Appl1 |
| 36 | 44 | 28.9 | 132 | 2 | US-08-472-281A-84 | Sequence 84, Appl1 |
| 37 | 44 | 28.9 | 132 | 2 | US-08-477-898B-84 | Sequence 84, Appl1 |
| 38 | 44 | 30.1 | 174 | 3 | US-09-122-443-6 | Sequence 6, Appl1 |
| 39 | 44 | 28.9 | 207 | 4 | US-09-336-536-50 | Sequence 50, Appl1 |
| 40 | 44 | 28.9 | 245 | 4 | US-09-336-536-48 | Sequence 48, Appl1 |
| 41 | 44 | 30.1 | 355 | 4 | US-09-134-001C-3460 | Sequence 3460, Ap |
| 42 | 44 | 30.1 | 759 | 4 | US-09-199-637A-170 | Sequence 170, App |
| 43 | 44 | 28.9 | 3075 | 2 | US-08-460-309-5 | Sequence 5, Appl1 |
| 44 | 44 | 28.9 | 3075 | 2 | US-08-125-077-5 | Sequence 5, Appl1 |
| 45 | 43 | 28.3 | 25 | 4 | US-09-336-536-62 | Sequence 62, Appl1 |

ALIGNMENTS

RESULT 1
US-09-479-309-2
; Sequence 2, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTSD0630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 239
; TYPE: PRT
; ORGANISM: human
US-09-479-309-2

Alignment Scores:

Pred. No.: 1,81e-16
Score: 152.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 3
Length: 239
Matches: 28
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x US-09-479-309-2 (1-239)

QY 1 ACTTCATTCCTCAGGTACAGACAGTGTGCTTCCTGCTTGGCTAACTTAAGAG 60
|||||
DB 13 ThierphehphatgryrarglncysleucysValProValValAlaAspPhelyslys 32
|||||
QY 61 CGGTGTTCTCAGATGATGAAGA 84
|||||
DB 33 ArgcypsheserGlutleuilearg 40

```
RESULT 2
US-09-134-001C-3138
; Sequence 3138, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3138
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3138

Alignment Scores:
Pred. No.: 0.0537 Length: 190
Score: 62.00 Matches: 12
Percent Similarity: 72.00% Conservative: 6
Best Local Similarity: 48.00% Mismatches: 7
Query Match: 40.79% Indels: 0
DB: 4 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x US-09-134-001C-3138 (1-190)

QY 10 TTAGAGTACAGACAGCTTTGCTGCTGCTGCTACTTAAAGCGCTGTTTC 69
DB 2 TTTGTTTATAspleuansleuNeuLeuValIleuYsIysArgCysPhe 21
QY 70 TCAGATTGATAGA 84
DB 22 SerLysMetIleArg 26

RESULT 3
US-09-082-059-3
; Sequence 3, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.
; APPLICANT: Devatajan, Prasad
; TITLE OF INVENTION: No. 6225086e1 Ankyrin Proteins and a Method for their Identificat
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082,059A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: a.a. seq., Ank-B, human kidney cDNA library
US-09-082-059-3

Alignment Scores:
Pred. No.: 5.93 Length: 69
Score: 49.00 Matches: 10
Percent Similarity: 46.67% Conservative: 4
Best Local Similarity: 33.33% Mismatches: 4
Query Match: 32.24% Indels: 12
DB: 4 Gaps: 1

US-09-939-293-1_COPY_56_139 (1-84) x US-09-082-059-3 (1-69)
```

```
QY 16 TACAGACAGTGTGCTGCTGCTGCTGCTACTT----- 54
DB 40 TTTATGTTTleIleCysValProTYrMetAlaIysPheValAlpHealIysSerHis 59
QY 55 -----AAGAACGGTGTTC 69
DB 60 AspProIleGluAlaIArgLeuAlrGysPhe 69

RESULT 4
US-08-466-103A-4
; Sequence 4, Application US/08466103A
; Patent No. 5856124
; GENERAL INFORMATION:
; APPLICANT: Reppert, Steven M.
; APPLICANT: Edisawa, Takashi
; TITLE OF INVENTION: HIGH-AFFINITY MELATONIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; SPOILER: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,103A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/319,887
; FILING DATE: 07-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,857
; FILING DATE: 17-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/250002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-466-103A-4

Alignment Scores:
Pred. No.: 7.16 Length: 366
Score: 49.00 Matches: 12
Percent Similarity: 88.24% Conservative: 3
Best Local Similarity: 70.59% Mismatches: 0
Query Match: 32.24% Indels: 2
DB: 2 Gaps: 2

US-09-939-293-1_COPY_56_139 (1-84) x US-08-466-103A-4 (1-366)

QY 11 TTAGTACAGACAGCTTTGCT---GCTGCTGCTGCTGCTGCTTAA 55
DB 156 SerGlyThrAsnSerLeuCysTyrValPheLeuIleIrrPThrLeuThrIleu 172

RESULT 5
US-08-896-365-7
```



```

Sequence 7, Application US/08896365
: Patent No. 5939264
:
GENERAL INFORMATION:
: APPLICANT: Rothschild, Max F.
: APPLICANT: Tuggle, Christopher K.
: APPLICANT: Messer, Lori A.
: APPLICANT: Tun-Ping, Yu
: TITLE OF INVENTION: GENES AND GENETIC MARKERS FOR IMPROVED
: TITLE OF INVENTION: REPRODUCTIVE TRAITS IN ANIMALS
: NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
: ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
: STREET: 801 Grand Avenue, Suite 3200
: CITY: Des Moines
: STATE: Iowa
: COUNTRY: USA
: ZIP: 50309
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/896,365
: FILING DATE: 18-JUL-1997
CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/022,180
: FILING DATE: 19-JUL-1996
ATTORNEY/AGENT INFORMATION:
: NAME: Nebel, Heidi S.
: REGISTRATION NUMBER: 37,719
: REFERENCE/DOCKET NUMBER: ISURP 021591
TELECOMMUNICATION INFORMATION:
: TELEPHONE: 515-288-3667
: TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 7:
:
SEQUENCE CHARACTERISTICS:
: LENGTH: 366 amino acids
: TYPE: amino acid
STRANDEDNESS: single
: TOPOLOGY: unknown
MOLECULE TYPE: CDNA
: HYPOTHETICAL: NO
ANTI-SENSE: NO
: ORIGINAL SOURCE:
ORGANISM: Ovis ammon aries
: CELL TYPE: Melanocyte
:
US-08-896-365-7
:
Alignment Scores:
Pred. NO.: 7.16 Length: 366
Score: 49.00 Matches: 12
Percent Similarity: 88.24% Conservative: 3
Best Local Similarity: 70.59% Mismatches: 0
Query Match: 32.24% Indels: 2
Gaps: 2
DB:
US-09-939-293-1_COPY_56_139 (1-84) x US-08-896-365-7 (1-366)
OY 11 TCAGGTACAGCAGTGTGG---GTCTTCCTGTGTGCG---CTACTTTA 55
|||||:::||||| |||||:::||| |||||
Db 156 SerGlyThrAsnSerLeucyStyValPheIleuIeTrpIleuThreu 172
RESULT 6
US-09-172-977-4
: Sequence 4, Application US/09172977
: Patent No. 5989863
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Guegler, Karl J.
: APPLICANT: Corley, Neil C.
: APPLICANT: Yue, Henry

```

```

US-09-939-293-1.COPY_56_139 (1-84) x US-09-172-977-4 (1-1839)
OY      16  TACAGACAGTGTGTTGTCCTGTTGGCGTAACATT----- 54
          |||||||:::|||||:::|||||:::|||||
Db       125  TYATGATGTTTTCTCCTGTTGGCGTAACATT----- 54
          |||||||:::|||||:::|||||:::|||||
OY      55  -----AAGAAGCCGCTGTTTC 69
          ::|||
Db       1305 AspProtleGuAlaArgLeuAtgCysPhe 1314

RESULT 7
US-09-247-155-109
; Sequence 109, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouquelercet, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET .021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 109
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-09-247-155-109

Alignment Scores:
Pred. NO.:           8.59             Length:         1839
Score:              49.00            Matches:         10
Percent Similarity: 46.67%           Conservative:    4
Best Local Similarity: 33.33%        Mismatches:     4
Query Match:        32.24%           Indels:         12
DB:                2                 Gaps:           1

US-09-939-293-1.COPY_56_139 (1-84) x US-09-172-977-4 (1-1839)
OY      16  TACAGACAGTGTGTTGTCCTGTTGGCGTAACATT----- 54
          |||||||:::|||||:::|||||:::|||||
Db       125  TYATGATGTTTTCTCCTGTTGGCGTAACATT----- 54
          |||||||:::|||||:::|||||:::|||||
OY      55  -----AAGAAGCCGCTGTTTC 69
          ::|||
Db       1305 AspProtleGuAlaArgLeuAtgCysPhe 1314

RESULT 7
US-09-247-155-109
; Sequence 109, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouquelercet, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET .021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 109
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
US-09-247-155-109

Alignment Scores:
Pred. NO.:           7.71             Length:         137
Score:              48.50            Matches:          9
Percent Similarity: 64.71%           Conservative:    2
Best Local Similarity: 52.94%        Mismatches:     3
Query Match:        31.91%           Indels:          3
DB:                4                 Gaps:            1

US-09-939-293-1.COPY_56_139 (1-84) x US-09-247-155-109 (1-137)

```

```

0Y      3 TTTCAATCTTCAGGTACAGACAGG-----TTTGCTGTCTCTGTGT 44
        |||:|||| ||| |||:|:| ||| ||| ||| ||| ||| ||| |||
Db      14 PheValLeuSerValThrIleIleIleCysPheThrCysSerCysCys 30

RESULT 8
; US-08-558-935-2
; Sequence 2, Application US/08558935
; Patent No. 6228637
; GENERAL INFORMATION:
; APPLICANT: Kasaoka, Keisuke
; APPLICANT: Kadoctani, Naoto
; APPLICANT: Kawata, Shigeru
; APPLICANT: Hayashi, Yumiko
; TITLE OF INVENTION: Recombinant Vector, Method for Giving
; TITLE OF INVENTION: Immunity Against PVY-T to Potato Plant,
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,935
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/139,157
; FILING DATE: 21-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-174P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: PVY-T coat protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Potato Virus Y-T
; US-08-558-935-2

Alignment Scores:
Pred. No.: 14.5 Length: 267
Score: 47.00 Matches: 8
Percent Similarity: 70.59% Conservative: 4
Best Local Similarity: 47.06% Mismatches: 5
Query Match: 32.19% Indels: 0
DB: 4 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x US-08-558-935-2 (1-267)
0Y      80 ATCAATTCTGAGAAACACCGCTTCTTAAGTTAGCCACAGAGAACACAC 30
        ::||| ||||| ::| ||| ::| |||:|||||||
Db      26 LeuAsnGlyGluGlyGlyAspLeuAsnValGlyThrSerGlyThrHis 42

RESULT 9
US-08-900-927-1

```

```

: Sequence1, Application US/08900927
: Patent No. 5840537
:
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Corley, Neil C.
: APPLICANT: Lal, Preeti
: TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Dr.
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/900,927
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PR-0350 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 593 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: MMLR20701
: CLONE: 475485
:
: US-08-900-927-1
:
: Alignment Scores:
: Pred. No.: 15.9 Length: 593
: Score: 47.00 Matches: 9
: Percent Similarity: 77.88 Conservative: 5
: Best Local Similarity: 50.00% Mismatches: 4
: Query Match: 30.92% Gaps: 0
: DB: 2
:
: US-09-939-293-1_COPY_56_139 (1-84) x US-08-900-927-1 (1-593)
:
: QY 20 GACAGTGTGTGTGTCCTGTGTGGCTAATTAGAGACGGGTGTTCTCAG 73
: ||| :: ||| |||||::: |||||::|||:::
: Db 402 AsphysileargValIleuLeuLeuTyrIleuLeuargGnsngIValserGlu 419
:
: RESULT 10
: US-08-900-927-3
: Sequence 3, Application US/08900927
: Patent No. 5840537
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Corley, Neil C.
: APPLICANT: Lal, Preeti
: TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Dr.
: CITY: Palo Alto
: STATE: CA

```

```

; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,927
; FILING DATE: Filed Herewith
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0350 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 642026
;
; US-08-900-927-3
;
; Alignment Scores:
; Pred. No.: 15.9 Length: 593
; Score: 47.00 Matches: 9
; Percent Similarity: 77.78% Conservative: 5
; Best Local Similarity: 50.00% Mismatches: 4
; Query Match: 30.92% Indels: 0
; DB: 2 Gaps: 0
;
; US-09-939-293-1_COPY_56_139 (1-84) x US-08-900-927-3 (1-593)
;
; QY 20 GACAGGTGTTGTCCTGTCGTCGCTAAGCTTAAGACGCGTGTCTCAG 73
; ||| ::: ||| |||||::: |||||:::|||||:::
; Db 402 AspylIleargValleuleuleuYrIleleuleuArgAnclYalserGIu 419
;
; RESULT 11
; US-08-900-927-4
; Sequence 4, Application US/08900927
; Patent No. 5840537
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,927
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0350 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1246217
;
; US-08-900-927-4
;
; Alignment Scores:
; Pred. No.: 15.9 Length: 593
; Score: 47.00 Matches: 9
; Percent Similarity: 77.78% Conservative: 5
; Best Local Similarity: 50.00% Mismatches: 4
; Query Match: 30.92% Indels: 0
; DB: 2 Gaps: 0
;
; US-09-939-293-1_COPY_56_139 (1-84) x US-08-900-927-4 (1-593)
;
; QY 20 GACAGGTGTTGTCCTGTCGTCGCTAAGCTTAAGACGCGTGTCTCAG 73
; ||| ::: ||| |||||::: |||||:::|||||:::
; Db 402 AspylIleargValleuleuleuYrIleleuleuArgAnclYalserGIu 419
;
; RESULT 12
; US-09-191-279-1
; Sequence 1, Application US/09191279
; Patent No. 5981192
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,279
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,927
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0350 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```

US-09-939-23-1_COPY_56_139 (1-84) x US-09-191-279-3 (1-593)
QY      20 GACAGTGTGTGTGTCCTGTCGTGAGCTTAACCTTAAGAACGGGTGTTCTCAG 73
      ||| ::| ||| |||||::|::| |||||::| |||||::|
Db      402 AsplysIleArGValIleuLeuLeuTyrlIleuLeuArGsnGlyValSerGlu 419

RESULT 14
US-09-191-279-4
; Sequence 4, Application US/09191279
; Patent No. 5981192
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,279
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,927
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0350 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 1246217
US-09-191-279-4

Alignment Scores:
Pred. No.:      15 9      length:      593
Percent:        47.00     Matches:       9
Score Similarity: 77.78%   Conservative: 5
Best Local Similarity: 50.00% Mismatches:  4
Query Match:    30.92%     Indels:      0
                        Gaps:      0

US-09-939-293-1_COPY_56_139 (1-84) x US-09-191-279-4 (1-593)
QY      20 GACAGTGTGTGTGTCCTGTCGTGAGCTTAACCTTAAGAACGGGTGTTCTCAG 73
      ||| ::| ||| |||||::|::| |||||::| |||||::|
Db      402 AsplysIleArGValIleuLeuLeuTyrlIleuLeuArGsnGlyValSerGlu 419

RESULT 15
US-09-334-476-1
; Sequence 1, Application US/09334476
; Patent No. 6162901
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.

```

APPLICANT: Lal, Preeti
TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/334,476
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/191,279
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0350 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 593 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: MMLR20701
CLONE: 475485
US-09-334-476-1

Alignment Scores:
Pred. No.: 15.9 Length: 593
Score: 47.00 Matches: 9
Percent Similarity: 77.78% Conservative: 5
Best Local Similarity: 50.00% Mismatches: 4
Query Match: 30.92% Indels: 0
DB: 4 Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x US-09-334-476-1 (1-593)

QY 20 GACAGTGTTCGTCTGCTTCCTGCTGCTTAAAGACGGGTTCACAG 73
||| :||| |||||:::| |||||:::| |||||:::|
Db 402 AspysileargValIleuLeuTyrlleuLeuargAsnglyValSerGlu 419

Search completed: February 16, 2003, 03:13:08
Job time : 10.7164 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 16, 2003, 01:20:19 ; Search time 28.209 Seconds

(without alignments)
793.581 Million cell updates/sec

Title: US-09-939-293-1_COPY_56_139

Perfect score: 152

Sequence: 1 actctatctctcaggtacagtaagtaaga 84

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09939293/runat_12022003_170352_8503/app_query.fasta_1.590
-DB=A.Geneseq_101002 -QFMT=fastan -SUFFIX=tag -MITMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINDEN=0 -MAXLEN=2000000000
-USER=us09939293 -CGEN=1.1.33 -runat_12022003_170352_8503 -NCPU=6 -ICPU=3
-NO_XLPUX -NO_MMAP -LARGEORDER -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.Geneseq_101002: *
1: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT: *
2: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT: *
3: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT: *
4: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT: *
5: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT: *
6: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT: *
7: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT: *
8: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT: *
9: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT: *
10: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT: *
11: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT: *
12: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT: *
13: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT: *
14: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT: *
15: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT: *
16: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT: *
17: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT: *
18: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT: *
19: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT: *
20: /SID2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT: *
21: /SID2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT: *
22: /SID2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT: *
23: /SID2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|---------------------|
| 1 | 152 | 100.0 | 239 | 21 | AA826210 | Human caspase acti |
| 2 | 152 | 100.0 | 239 | 21 | AAU78447 | Inhibitor of apopt |
| 3 | 147 | 96.7 | 227 | 21 | AA854139 | Human pancreatic c |
| 4 | 62 | 40.8 | 190 | 23 | ABP38293 | Staphylococcus epi |
| 5 | 54.5 | 35.9 | 687 | 23 | ABB62975 | Herbicideally activ |
| 6 | 51.5 | 35.3 | 1099 | 22 | ABB68903 | Drosophila melanog |
| 7 | 50.5 | 34.6 | 4260 | 22 | AA823830 | Human ESR encoded |
| 8 | 50.5 | 34.6 | 4854 | 22 | ABG06027 | Novel human diagno |
| 9 | 50 | 32.9 | 122 | 21 | AA820017 | Arabidopsis thalia |
| 10 | 50 | 32.9 | 141 | 21 | AA620016 | Arabidopsis thalia |
| 11 | 50 | 32.9 | 163 | 21 | AA620015 | Arabidopsis thalia |
| 12 | 50 | 34.2 | 299 | 22 | ABB65597 | Drosophila melanog |
| 13 | 49.5 | 33.9 | 763 | 23 | ABB50029 | Listeria monocytog |
| 14 | 49 | 32.2 | 69 | 22 | AA862430 | Human kidney ankyr |
| 15 | 49 | 32.2 | 242 | 22 | AA836606 | Human FLEKHT-28 pr |
| 16 | 49 | 32.2 | 366 | 17 | AA88410 | High-effinity mela |
| 17 | 49 | 32.2 | 495 | 22 | ABB61724 | Drosophila melanog |
| 18 | 49 | 32.2 | 545 | 23 | ABB89989 | Human polypeptide |
| 19 | 49 | 32.2 | 632 | 23 | ABB93993 | Herbicideally activ |
| 20 | 49 | 32.2 | 636 | 22 | AA894352 | Human protein sequ |
| 21 | 49 | 32.2 | 644 | 22 | ABG06230 | Novel human diagno |
| 22 | 49 | 32.2 | 747 | 22 | AA840070 | Human polypeptide |
| 23 | 49 | 32.2 | 747 | 22 | AA893124 | Human protein sequ |
| 24 | 49 | 32.2 | 1872 | 22 | AA879160 | Human protein seq |
| 25 | 48.5 | 31.9 | 118 | 23 | ABP61439 | Human NF-kB activa |
| 26 | 48.5 | 31.9 | 137 | 20 | AAV59678 | Secreted protein 1 |
| 27 | 48.5 | 31.9 | 137 | 21 | AAV84606 | A human small prol |
| 28 | 48.5 | 31.9 | 137 | 22 | AA839215 | Human polypeptide |
| 29 | 48.5 | 31.9 | 137 | 22 | AA831675 | Amino acid sequenc |
| 30 | 48.5 | 31.9 | 137 | 23 | ABP61438 | Human NF-kB activa |
| 31 | 48.5 | 33.2 | 138 | 22 | AAU87606 | Novel central nerv |
| 32 | 48.5 | 33.2 | 138 | 22 | AAE09642 | Human gene 4 encod |
| 33 | 48.5 | 31.9 | 155 | 19 | AAW75148 | Human secreted pro |
| 34 | 48.5 | 31.9 | 155 | 19 | AAW75087 | Human secreted pro |
| 35 | 48.5 | 31.9 | 179 | 22 | AA841001 | Human polypeptide |
| 36 | 48.5 | 33.2 | 220 | 22 | AAU87317 | Human central nerv |
| 37 | 48.5 | 31.9 | 240 | 22 | AAE03325 | Human gene 6 encod |
| 38 | 48.5 | 31.9 | 240 | 22 | AAE03325 | Human gene 6 encod |
| 39 | 48.5 | 31.9 | 240 | 22 | AA888389 | Human membrane or |
| 40 | 48.5 | 31.9 | 240 | 23 | ABG64408 | Human albumin fusi |
| 41 | 48.5 | 31.9 | 240 | 23 | ABG64409 | Human albumin fusi |
| 42 | 48 | 31.6 | 60 | 22 | AAO04087 | Human polypeptide |
| 43 | 48 | 32.9 | 110 | 22 | AAU50894 | Proionibacterium |
| 44 | 48 | 32.9 | 191 | 22 | AAU29508 | Human G protein-co |
| 45 | 48 | 32.9 | 191 | 23 | ABG60796 | Novel G protein co |

ALIGNMENTS

| | | |
|----------|-------------|--|
| RESULT 1 | AA826210 | standard; Protein: 239 AA. |
| ID | AA826210 | standard; Protein: 239 AA. |
| XX | AA826210; | |
| AC | 23-FEB-2001 | (first entry) |
| XX | | |
| DT | | |
| XX | | |
| DE | | Human caspase activator Smac. |
| XX | | |
| XX | | Human: caspase activator; Smac; Apoptosis; cancer; autoimmune disease; |
| KW | | neurodegenerative disease; mitochondria. |
| XX | | |
| OS | | Homo sapiens. |
| XX | | |
| PN | | US6110691-A. |
| XX | | |
| PD | | 29-AUG-2000. |
| XX | | |

PF 06-JAN-2000; 2000US-0479309.
 XX
 PR 06-JAN-2000; 2000US-0479309.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA
 XX Wang X, Du C;
 PI
 XX WPI: 2000-586350/55.
 DR N-PSDB; AAA94860.
 XX
 PT Novel caspase regulatory polypeptides useful for screening binding
 agents specific for the polypeptides which are useful for diagnosis and
 also for treating apoptosis associated diseases -
 PS Claim 1; column 23-24; 16pp; English.
 XX
 CC The present sequence is the human Smac protein. Its coding sequence
 CC was isolated by purifying the protein and searching a HeLa cell cDNA
 CC library for sequences which bound to probes based upon it. Smac is a
 CC mitochondrial protein which is released into the cytosol during
 CC apoptosis, and acts as a caspase-3 activator. The protein and its coding
 CC sequence can be used to modulate the expression and function of caspases
 CC and their activators, and also can be used as drug targets and regulators
 CC to promote or inhibit apoptosis in the treatment of cancer and autoimmune
 CC and neurodegenerative diseases.
 XX
 SQ Sequence 239 AA:
 Alignment Scores:
 Pred. No.: 3.12e-14 Length: 239
 Score: 152.00 Matches: 28
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-939-293-1_COPY_56_139 (1-84) x AAB26210 (1-239)
 QY 1 ACTTCATCTTCAGGTACAGAGCTGTTCCGTGCTGGCTAACTTAAGANG 60
 DB 13 ThnSerPhePheArgTyrArgGlnCysLeuValProValValAlaAsnPhenylslys 32
 QY 61 CGGTGTTCTCAGAAATGTATAGA 84
 DB 33 ArgCysPheSerGlnLeuLeuArg 40
 RESULT 2
 AAU78447
 ID AAU78447 standard; Protein; 239 AA.
 AC AAU78447;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Inhibitor of apoptosis (IAP) protein Smac.
 XX
 KM Human; inhibitor of apoptosis; IAP; Smac; apoptosis; BID; BIR1; BIR2;
 KW Bcl2 interacting domain; caspase; BIR domain; BIR3; gene therapy;
 KM neoplastic cell; tumour.
 XX
 OS Homo sapiens.
 OS
 PN WO200216418-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 24-AUG-2001; 2001WO-US26492.
 XX
 PR 24-AUG-2000; 2000US-227735P.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX

PI Alnemrl ES;
 XX
 DR WPI: 2002-304115/34.
 DR N-PSDB; ABR15451.
 XX
 PT Novel Smac peptides and polynucleotides encoding the peptides, useful
 PT for stimulating apoptosis in neoplastic or tumour cell which
 PT overexpresses inhibitor of caspase, and for identifying apoptosis
 PT modulating compounds -
 PS Claim 36; Page 73-74; 78pp; English.
 XX
 CC The invention relates to an isolated Smac peptide or polypeptide (I)
 CC and an isolated nucleic acid (II) encoding (I). Also described is a
 CC method of identifying a compound that inhibits apoptosis, comprising:
 CC (a) separately contacting several cell populations expressing a
 CC cytosolic Smac (a Smac isoform that begins with MKSDPRF sequence,
 CC replacing the mitochondrial targeting sequence (residues 1-35 of (I)),
 CC and residues 56-60 of (I)) and an inhibitor of BID (Bcl2 interacting
 CC domain) with a compound to be tested for apoptotic inhibiting activity;
 CC (b) incubating the cell populations with a direct stimulus of the cell
 CC death pathway; and (c) measuring the specific apoptotic activity of the
 CC cell populations, where inhibition of the specific apoptotic activity is
 CC indicative that the compound is an inhibitor of apoptosis. (I) and (II)
 CC are useful for inducing apoptosis in a cell. The Smac polypeptide and
 CC polynucleotide are useful for stimulating apoptosis in a neoplastic or
 CC tumour cell which overexpresses an inhibitor of caspase, where the
 CC inhibitor inhibits activation or activity of caspase-3, caspase-7 or
 CC caspase-9. Preferably, the cell overexpresses at least a portion of IAP.
 CC (I) is useful for identifying an inhibitor or enhancer of a caspase-
 CC mediated apoptosis which involves contacting a cell transformed or
 CC transfected with a vector expressing (I) with a candidate inhibitor or
 CC candidate enhancer; and detecting cell viability, where an increase in
 CC cell viability indicates the presence of an inhibitor and a decrease in
 CC cell viability indicates the presence of an enhancer. Optionally, the
 CC method involves detecting the presence of large and small caspase
 CC subunits after contacting cell transformed with the vector expressing
 CC (I), with the candidate compound. A decrease in processing indicates the
 CC presence of an inhibitor and an increase in processing indicates the
 CC presence of an enhancer. Preferably, the large and small subunits of
 CC caspase-3, caspase-7 or caspase-9 are detected. (I) is also useful for
 CC identifying a compound that inhibits Smac binding to Smac-binding
 CC molecule (a portion of IAP e.g. a BIR domain such as BIR1, BIR2 or BIR3,
 CC or a full-length IAP). (II) is useful in gene therapy techniques. The
 CC present sequence represents the amino acid sequence of Smac protein.
 XX
 SQ Sequence 239 AA:
 Alignment Scores:
 Pred. No.: 3.12e-14 Length: 239
 Score: 152.00 Matches: 28
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 23 Gaps: 0
 US-09-939-293-1_COPY_56_139 (1-84) x AAU78447 (1-239)
 QY 1 ACTTCATCTTCAGGTACAGAGCTGTTCCGTGCTGGCTAACTTAAGANG 60
 DB 13 ThnSerPhePheArgTyrArgGlnCysLeuValProValValAlaAsnPhenylslys 32
 QY 61 CGGTGTTCTCAGAAATGTATAGA 84
 DB 33 ArgCysPheSerGlnLeuLeuArg 40
 RESULT 3
 AAB54139
 ID AAB54139 standard; Protein; 227 AA.
 XX
 AC AAB54139;
 XX
 DT 09-MAR-2001 (first entry)
 XX

DE Human pancreatic cancer antigen protein sequence SEQ ID NO:591.

XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW detection; diagnosis; identification; cytostatic; neuroprotective;
 KW neoplastic; immunomodulatory; relaxant; contraceptive; gynaecological;
 KW antiinflammatory; cardiact; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative.

XX Homo sapiens.

OS NO200005320-A1.

PN 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05989.

PR 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM;

PI WPI: 2000-579444/54.

DR N-PSDB; AAC98904.

PT New nucleic acid that is a pancreatic cancer antigen for preventing,
 PT treating, or ameliorating a medical condition, particular pancreatic
 PT cancer, or for use in assays for diagnosing a pathological condition -
 XX

PS Claim 11: Page 1027-1028; 1379pp; English.

XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,
 CC neuroprotective, neoplastic, immunomodulatory, relaxant, contraceptive,
 CC gynaecological, cardiact and antiinflammatory activities, and can be used
 CC in gene therapy. The polynucleotide and proteins can be used for
 CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Abonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in chromosome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect and target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.

XX

XX SQ Sequence 227 AA:

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 1.73e-13 | Length: | 227 |
| Score: | 147.00 | Matches: | 27 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 96.71% | Indels: | 0 |
| DB: | 21 | Gaps: | 0 |

US-09-939-293-1_COPY_56_139 (1-84) x AAB54139 (1-227)

0Y 4 TCATCTTCAGGACAGACAGTGTGTGTCTGTGTGGCTTAACATTTAAGACGG 63

2 Serpethenargrtyrgrtncysleucysvalpvovalvalaiaasnphylslysrq 21

```

OY 64 TGTTCTCAGATTGATTAAGA 84
      |||||||||||||||||||
Db 22 CysPheSerGlnLeuIleArg 28

RESULT 4
ID ABP38293 standard; Protein; 190 AA.
XX
XX ABP38293:
XX
XX 24-JUL-2002 (first entry)
XX
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3138.
XX
KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection
XX antibacterial; gene therapy.
XX
OS Staphylococcus epidermidis.
XX
XX US6380370-B1.
XX
XX 30-APR-2002.
XX
XX 13-AUG-1998; 98US-0134001.
XX
XX 14-AUG-1997; 97US-055779P.
XX PR 08-NOV-1997; 97US-064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
XX
XX WPI; 2002-381255/41.
XX DR N-PSDB; ABN90838.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
XX polypeptide, useful for diagnosing and treating bacterial infections -
XX
XX Disclosure; SEQ ID 3138; 267pp; English.
XX
CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life
CC cycle or inhibit S. epidermidis infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.
XX
XX
SQ Sequence 190 AA;
XX
XX
XX Alignment Scores:
XX Pred. NO.: 0.847 Length: 190
XX Score: 62.00 Matches: 12
XX Percent Similarity: 72.00% Conservatve: 6
XX Best Local Similarity: 48.00% Mismatches: 7
XX Query Match: 40.79% Indels: 0
XX DB: 23 Gaps: 0
XX
US-09-939-293-1_COPY_56_139 (1-84) x ABP38293 (1-190)
OY 10 TTTCAGGTACACAGCTGTTGTGTCGTCCTCGTTGTGCTACTTAAAGAAGCGGTGTTCC 69
      ::::||||| ||| ::::||||| ||| |||||||||||||
Db 2 TyriIntYirspLeuAnaSenLeuLeuProLeuValYlsIleLeuYlsYsArgCysPhe 21
OY 70 TCAGATTGATTAAGA 84
      |||:|||||:|||||
Db 22 SerYsMetIleArg 26

RESULT 5

```

ABB92975
ID ABB92975 standard; Protein; 687 AA.
XX
AC ABB92975;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2186.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidner M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 2186; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 687 AA;
XX
Alignment Scores:
Pred. No.: 12 6
Score: 54.50 Length: 687
Percent Similarity: 65.228 Matches: 10
Best Local Similarity: 43.488 Conservative: 5
Query Match: 35.868 Mismatches: 7
Indels: 1
Gaps: 1
US-09-939-293-1_COPY_56_139 (1-84) x ABB92975 (1-687)
OY 1 ACTTCATCTTCAGTACAGACAG---TGTGTGCTGCTTCCTGTCGCTAACTTAAAG 57
DB 240 ThrasntrphelysTylnsGlnThncysLysProthrProtyrValGlnAsnPhnasp 239
OY 58 AACGGGTGT 66
DB 260 ArgAlacys 262
RESULT 6
ABB68903
ID ABB68903 standard; Protein; 1099 AA.
XX
AC ABB68903;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 33501.

XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL13006.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 33501; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB5737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 1099 AA;
XX
Alignment Scores:
Pred. No.: 37 1
Score: 51.50 Length: 1099
Percent Similarity: 72.008 Matches: 13
Best Local Similarity: 52.008 Conservative: 5
Query Match: 35.278 Mismatches: 4
Indels: 3
Gaps: 1
US-09-939-293-1_COPY_56_139 (1-84) x ABB68903 (1-1099)
OY 84 TCTTATCAATTCAGAACACCGCTTCCTTAAGTACCAACAAGAACACACAACA 25
DB 789 SerTyrIleValGlnThrProLeuProLysIleSer-----AsnThrArgThr 805
OY 24 CTGTCTGTACCTGAA 10
DB 806 LeuAsnValLeuGln 810
RESULT 7
AAM23830
ID AAM23830 standard; Protein; 4260 AA.
XX
AC AAM23830;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST encoded protein SEQ ID NO: 1355.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;

KM biodiversity; gene therapy; nutrition.
 XX
 OS Homo sapiens.
 XX
 XX WO200154477-A2.
 XX
 XX 02-AUG-2001.
 PD
 XX 25-JAN-2001; 2001WO-US02687.
 PF
 XX 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 XX
 DR WPI: 2001-476164/51.
 DR N-PSDB: AAH98489.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -
 XX
 PS Claim 20; Page 955-964; 1275pp: English.
 XX
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.
 XX
 SQ Sequence 4260 AA;
 Alignment Scores:
 Pred. No.: 59.6 Length: 4260
 Score: 50.50 Matches: 12
 Percent Similarity: 60.71% Conservative: 5
 Best Local Similarity: 42.86% Mismatches: 10
 Query Match: 34.59% Indels: 1
 DB: 22 Gaps: 1
 US-09-939-293-1_COPY_56_139 (1-84) x AAM23830 (1-4260)
 QY 83 CTTATCAATTCTGAGAACACCGCTTCTTA---AGTTAGCCACACAGAACACACAA 27
 DB 3725 LeuLeuGluSerIleGlnIleValPheIleLysValAlaValAsnSerGlyGlyLys 3744
 QY 26 CACTGCTGTACTGTAAGATGAA 3
 DB 3745 HisCysLeuAlaLeuSerSerGlu 3752
 RESULT 8
 ABG06027
 ID ABG06027 standard; Protein: 4854 AA.
 XX
 XX ABG06027;
 AC
 XX
 XX 13-FEB-2002 (first entry)
 DT
 XX
 XX Novel human diagnostic protein #6018.
 DE
 XX
 XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX
 XX WO200175067-A2.

XX
 PD 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS70214.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 36386; 103pp: English.
 XX
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 4854 AA;
 Alignment Scores:
 Pred. No.: 60.4 Length: 4854
 Score: 50.50 Matches: 12
 Percent Similarity: 60.71% Conservative: 5
 Best Local Similarity: 42.86% Mismatches: 10
 Query Match: 34.59% Indels: 1
 DB: 22 Gaps: 1
 US-09-939-293-1_COPY_56_139 (1-84) x ABG06027 (1-4854)
 QY 83 CTTATCAATTCTGAGAACACCGCTTCTTA---AGTTAGCCACACAGAACACACAA 27
 DB 4050 LeuLeuGluSerIleGlnIleValPheIleLysValAlaValAsnSerGlyGlyLys 4069
 QY 26 CACTGCTGTACTGTAAGATGAA 3
 DB 4070 HisCysLeuAlaLeuSerSerGlu 4077
 RESULT 9
 AAG20017
 ID AAG20017 standard; Protein: 122 AA.
 XX
 XX AAG20017;
 AC
 XX
 XX 17-OCT-2000 (first entry)
 DT
 XX
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 22040.
 DE

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hydralisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139765.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145102.
PR 22-JUL-1999; 99US-0145119.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-01452913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

```
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160778.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161892.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Alignment Scores:

```
Pred. No.: 50.4
Score: 50.00
Percent Similarity: 84.628
Best Local Similarity: 61.548
Query Match: 32.898
DB: 21
Matches: 122
Conservative: 8
Mismatch: 2
Indels: 0
Gaps: 0
```

US-09-939-293-1_COPY_56_139 (1-84) x AAG20017 (1-122)

```
OY 8 TCTTCAGGTACAGACAGCTGTTGTGTCCTGTTGTGG 46
|||||:::|||||
Db 99 SerSerSerSerSerSerValCysValPheMetValTrp 111
```

RESULT 10

AAG20016 AAG20016 standard; Protein; 141 AA.

AC AAG20016;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 22039.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000, 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129645.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134421.

PR 18-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0143390.
PR 08-JUL-1999; 99US-0144803.
PR 09-JUL-1999; 99US-0144920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 51.1
Score: 50.00
Percent Similarity: 84.62%
Best Local Similarity: 61.54%
Query Match: 32.89%
DB: 21
Gaps: 0

US-09-939-293-1_COPY_56_139 (1-84) x AAG20016 (1-141)

OY 8 TCTTCAGGTACACAGTGTGTGTCCTGCTGCG 46

Db 118 SerSerSerSerSerValCysValPheMetValIrrp 130

RESULT 11

AAG20015
ID AAG20015 standard; Protein: 163 AA.

AC AAG20015;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment seq ID NO: 22038.

KM Protein identification; signal transduction pathway; metabolic pathway;

KW hydridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

```
XX Arabidopsis thaliana.
OS
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
```

```

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Alignment Scores:
Pred. No.: 51.8 Length: 163
Score: 50.00 Matches: 8
Percent Similarity: 84.62% Conservative: 3
Best Local Similarity: 61.54% Mismatches: 2
Query Match: 32.89% Indels: 0
DB: 21 Gaps: 0

```

US-09-939-293-1_COPY_56_139 (1-84) x AAG20015 (1-163)

```

OY 8 TCTTCAGGTACAGACGTTGTGTGTCCTGTTGTG 46
    ||||| :||| ||||| ||||| :|||
DB 140 SerSerSerSerSerSerValCysValPheMetValTrp 152

```

RESULT 12

ID ABB65597 standard; Protein: 299 AA.

AC ABB65597;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 23583.

KW Drosophila: developmental biology; cell signalling; insecticide;

KW pharmaceutical.

OS Drosophila melanogaster.

XX

```

PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PERE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX
DR N-PSDB; ABL09700.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
XX Disclosure: SEQ ID NO 23583; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 299 AA:
SQ

```

```

Alignment Scores:
Pred. No.: 54.9 Length: 299
Score: 50.00 Matches: 9
Percent Similarity: 66.67% Conservative: 5
Best Local Similarity: 42.86% Mismatches: 7
Query Match: 34.25% Indels: 0
DB: 22 Gaps: 0

```

US-09-939-293-1_COPY_56_139 (1-84) x ABB65597 (1-299)

```

OY 83 CTATCAATTCGAGAAACACCGCTTTAAAGTTAGCCACACGACACAAAC 24
    ||||| :||| ||||| ||||| :|||
DB 204 LeuValGlnAsnGluLysHisSerGlnHisLysLeuIleThrSerGlyIleTyrAlaTyr 223

```

```

OY 23 TGT 21
    |||
DB 224 Cys 224

```

RESULT 13

ID ABB50029 standard; Protein: 763 AA.

XX ABB50029;

XX 05-FEB-2002 (first entry)

XX Listeria monocytogenes protein #2733.

KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

KW vitamin B12; bacterial infection; disease.

OS Listeria monocytogenes.

XX WO200177335-A2.

XX 18-OCT-2001.

XX

PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangoul L, Couve E, Rusniok C, Fahl H, Dehoux P,
PI Dussurget O, Chetouani F, Nedjar H, Glaeser P, Kunst F, Cossart P,
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Roland JA,
PI Dominguez-Bernal G, Garrido-Garcia P, Rietz-Martinez A, Amend A,
PI Chakraborty T, Doman E, Hain T, Berche P, Chardit A, Durant L,
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
PI Madueno E, de Pablos B, Wehland J, Kaerst U, Entian K, Hauf J,
PI Rose M, Voss H;
XX
XX WPI; 2002-010914/01.
XX
XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
PT and prevention of *Listeria* and related bacterial infections, and
PT related polypeptides -
XX
XX Claim 6; SEQ ID No 2734; 192pp; French.
XX
XX The present invention relates to the genome sequence of *Listeria*
CC *monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in *L.*
CC *monocytogenes* and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of *L. monocytogenes* and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by *L.*
CC *monocytogenes* and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 763 AA;
XX
XX Alignment Scores:
XX Pred. No.: 71.3 Length: 763
XX Score: 49.50 Matches: 12
XX Percent Similarity: 64.00% Conservative: 4
XX Best Local Similarity: 48.00% Mismatches: 8
XX Query Match: 33.90% Indels: 1
XX DB: 23 Gaps: 1
XX
XX US-09-939-293-1_COPY_56_139 (1-84) x ABB50029 (1-763)
XX
XX QY 80 ATCAATTCGAGAAACGCGTCTTAAAGTTAGCCACAGACACAAACACTGT 21
XX ||| :||||:||||| :||| ||| ||| ||| |||
XX Db 536 TlEThrAsmGlnYsTyRArgSerIlleGlyLeuGlyThrphGlyTrpHis---HisLeu 554
XX
XX QY 20 CTGTACTCTGAAGAAT 6
XX ||| ||| ||| ||| |||
XX Db 555 LeuAlaLeuLysAsn 559
XX
XX RESULT 14
XX AAB62430
XX ID AAB62430 standard; Protein; 69 AA.
XX
XX AC AAB62430;
XX
XX 09-JUL-2001 (first entry)
XX
XX DE Human kidney ankrylin-B protein fragment.
XX
XX Ankyrin G119; AnkG119; kidney; human; spectrin binding domain;

KW ankrylin-B; integral membrane protein; secretory protein.
XX
XX OS Homo sapiens.
XX
XX PN US6225086-B1.
XX
XX PD 01-MAY-2001.
XX
XX PF 21-MAY-1998; 98US-0082059.
XX
XX PR 21-MAY-1997; 97US-0047356.
XX
XX (UYVA) UNIV YALE.
XX
XX PI Morrow JS, Devorajan P;
XX
XX WPI; 2001-315683/33.
XX
XX DR Novel isolated nucleic acid encoding ankrylin G119 protein, useful for
PT producing recombinant ankrylin G119 protein which participates in
PT trafficking of secretory proteins between endoplasmic reticulum and
PT compartments -
XX
XX Example 1; Fig 1B; 40pp; English.
XX
XX The invention relates to a human ankrylin G119 (AnkG119) protein. The
CC AnkG119 has a molecular weight of 116 kDa and comprises a region
CC consisting of 13 repeats of a 33 residue structure and further comprises
CC a spectrin binding domain and 5 kDa regulatory domain. The AnkG119
CC nucleic acid is useful for producing the protein by standard recombinant
CC techniques. The AnkG119 protein participates in the trafficking of
CC integral membrane proteins and secretory proteins between the endoplasmic
CC reticulum and other membrane compartments. The present sequence
CC represents a human kidney ankrylin-B protein fragment.
XX
XX SQ Sequence 69 AA;
XX
XX Alignment Scores:
XX Pred. No.: 67.3 Length: 69
XX Score: 49.00 Matches: 10
XX Percent Similarity: 46.67% Conservative: 4
XX Best Local Similarity: 33.33% Mismatches: 4
XX Query Match: 32.24% Indels: 12
XX DB: 22 Gaps: 1
XX
XX US-09-939-293-1_COPY_56_139 (1-84) x AAB62430 (1-69)
XX
XX QY 16 TACAGACAGGTGTGTGCTGTTGCTGTTGCGTAACTTT----- 54
XX |||||: :||| ||| ||| :||| |||
XX Db 40 TyRArgGlnIleIleCysValProTyrMetAlaLysPheValPheAlaLysSerHis 59
XX
XX QY 55 -----AAGAAGCGGTGTTTC 69
XX :||| ||| |||
XX Db 60 AspProIleGlnAlaArgLeuArgCysPhe 69
XX
XX RESULT 15
XX AAB36606
XX ID AAB36606 standard; Protein; 242 AA.
XX
XX AC AAB36606;
XX
XX 09-MAR-2001 (first entry)
XX
XX DE Human FLEXHT-28 protein sequence SEQ ID NO:28.
XX
XX Human FLEXHT: full-length molecules expressed in human tissue;
KW diagnosis; gene expression; genetic linkage; genetic variability;
KW antineutemic; anticoagulant; antiarteriosclerotic; immunomodulatory;
KW cytoskeletal; anticonvulsant; antiinflammatory; hepatotropic; antidiabetic;
KW anti-gout; antithyroid; neuroprotective; antiallergic; osteopathic;
KW antipsoriatic; antirheumatic; antilucer; gene therapy; anaemia; gout;
KW epilepsy; arteriosclerosis; atherosclerosis; developmental disorder;
KW cancer; immunological disorder; asthma; bronchitis; cirrhosis;

KW Crohn's disease; diabetes mellitus; Grave's disease; multiple sclerosis;
 KW osteoarthritis; pancreatitis; rheumatoid arthritis; psoriasis;
 KW ulcerative colitis.
 OS Homo sapiens.
 XX
 XX WO200070047-A2.
 XX
 XX PD
 XX 23-NOV-2000.
 XX
 XX PF 12-MAY-2000; 2000MO-US13299.
 XX
 XX PR 14-MAY-1999; 99US-0311894.
 XX PR 14-MAY-1999; 99US-0311937.
 XX PR 14-MAY-1999; 99US-0311940.
 XX
 XX PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX PI Yue H, Tang YT, Lal P, Reddy R, Batra S, Baughn MR, Yang J;
 XX PI Azimzal Y, Lu DM, Au-Young J, Shih LL;
 XX
 XX DR N-PSDB; AAC88097.
 XX
 XX DR WPI: 2001-016234/02.
 XX
 XX PS Claim 1; Page 118-119; 168pp; English.
 CC AAC88070 to AAC88124 encode the 55 FLEXHT (full-length molecules
 CC expressed in human tissues) proteins given in AAB36579 to AAB36633. The
 CC present invention describes an isolated polypeptide (A) comprising an
 CC amino acid sequence selected from one of 55 amino acid sequences 42-876
 CC residues in length, corresponding to FLEXHT-1 to FLEXHT-55, a 90 %
 CC identical sequence, and a biologically active or immunogenic fragment of
 CC the sequence. The FLEXHT proteins can have antianaemic, anticonvulsant,
 CC antiarteriosclerotic, immunomodulatory, cytosolic, antisthmatic,
 CC antiinflammatory, hepatotropic, antidiabetic, anti-gout, antithyroid,
 CC neuroprotective, antiarthritis, osteopathic, antipsoriatic, antiulcer
 CC and antirheumatic activities, and can be used in gene therapy. The
 CC polynucleotide sequences can be used to express the protein sequences.
 CC Pharmaceutical compositions comprising FLEXHT can be used to treat
 CC diseases or conditions associated with altered expression of functional
 CC FLEXHT. The proteins and polynucleotides can be used to diagnose and
 CC treat disorders including anaemia, epilepsy, arteriosclerosis,
 CC atherosclerosis, developmental disorders, cancers, and immunological
 CC disorders such as asthma, bronchitis, cirrhosis, Crohn's disease,
 CC diabetes mellitus, gout, Grave's disease, multiple sclerosis,
 CC osteoarthritis, pancreatitis, psoriasis, rheumatoid arthritis, and
 CC ulcerative colitis.
 CC
 CC SQ Sequence 242 AA;
 Alignment Scores:
 Pred. No.: 75.9 Length: 242
 Score: 49.00 Matches: 8
 Percent Similarity: 66.67% Conservative: 2
 Best Local Similarity: 53.33% Mismatches: 5
 Query Match: 32.24% Indels: 0
 DB: 22 Gaps: 0
 US-09-939-293-1_COPY_56_139 (1-84) x AAB36606 (1-242)
 QY 22 CAGTGTTCGTGTCCTGTCGCTACTTATAGAACGGGCT 66
 |||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 168 GincysmeCProva1Prova1lleuasnPhaaspAlaGluCys 182